JIANMO FANGFALUN

SHUXUE

数学建模方法论

主 编 段克峰



兰州大学出版社

SHUXUE JIANMO FANGFALUN

策划编辑/梁建萍　责任编辑/郝可伟　封面设计/杨佩哲

ISBN 978-7-311-04232-5



定价：25.00元

陇东学院教材学术著作基金资助出版

JIANMO FANGFALUN

SHUXUE

数学建模方法论

主　编　段克峰
副主编　许万银
编　者　赵慧霞　杨大勇



兰州大学出版社

图书在版编目(CIP)数据

数学建模方法论／段克峰主编．—兰州：兰州大学出版社，2013.8

ISBN 978-7-311-04232-5

Ⅰ.①数… Ⅱ.①段… Ⅲ.①数学模型—研究 Ⅳ.①022

中国版本图书馆 CIP 数据核字(2013)第 201278 号

策划编辑 梁建萍
责任编辑 郝可伟
封面设计 杨佩哲

书　　名 数学建模方法论
作　　者 段克峰 主编
出版发行 兰州大学出版社 (地址:兰州市天水南路 222 号 730000)
电　　话 0931-8912613(总编办公室) 0931-8617156(营销中心)
　　　　 0931-8914298(读者服务部)
网　　址 http://www.onbook.com.cn
电子信箱 press@lzu.edu.cn
印　　刷 兰州万易印务有限责任公司
开　　本 710 mm×1020 mm 1/16
印　　张 11.25
字　　数 197 千
版　　次 2013 年 8 月第 1 版
印　　次 2013 年 8 月第 1 次印刷
书　　号 ISBN 978-7-311-04232-5
定　　价 25.00 元

前 言

当今社会日趋数学化，正如数学家笛卡儿所言：“一切问题都可以化成数学问题。”高技术实则为一种数学技术。数学技术是指把现实的问题转述成一个相应的数学模型，且用计算机加以解决或用数学理论定性、定量地加以研究，得出那个现实问题的定量结论或重要性质。所谓数学模型，乃是现实世界当中某一类特殊的运动变化过程或关系或结构的一种模拟性的数学结构，是对现实模型的理想化，是一种科学的抽象。伟大的科学家约翰·冯·诺依曼(John von Neumann)云：“科学的目的不只是解释现象，科学的主要目的是建立数学模型。”但是，如何建立和解答纷纭复杂的实际问题的数学模型呢?

数学建模与其说是一门技术，不如说是一门艺术。与一门技术大致有章可循不同，艺术在某种意义下无法归纳出若干条普遍适用的准则或方法。尽管如此， 数学建模还是有一些比较常用的基本的技术手段和方法。

考虑到数学建模是一门内容活泼、信息量大、涉及数学多个分支的课程，在实际应用中，不但需要掌握相关数学分支的原理和方法，还要有实际问题的背景知识，无论是对从事数学建模教学的老师还是学生都是挑战。特别是培训参加数学建模竞赛学生的大学老师，由于教学科研任务繁重，希望能找到一本能在较少的课时内使学生掌握数学建模的主要内容和方法的教材；而参加数学建模竞赛的同学不但有正常的课程学习任务，还由于专业分布广泛，对有关的数学知识不可能也没必要钻得很深，也希望找到一本介绍数学知识适可而止，以实用为度的参考书。

鉴于以上原因，编者在多年从事 “数学模型”、“数学实验” 课程的教学和组织指导数学建模竞赛工作的基础上，阅读了大量的相关资料，总结了多年从事数学建模教学和研究工作的经验，编著了本书。

本书以众多典型的具体问题的分析解答，向读者介绍了建立数学模型的思路与方法、技巧。通过一些典型数学建模案例的分析，详细介绍了几种数学建模方法。

在建模过程中，不得不把那些对我们关心的问题影响甚微的因素忽略掉，不然，所建模型因为数学结构太复杂而失去数学上的可解性；但也不能把足够相关的因素忽略掉，不然所建模型因为不能足够准确地反映实际情况而失去可靠性。可解性与可靠性同时最佳是很罕见的，一般总是在可解性的前提下，力争有满意的可靠性。

同一个实际问题，可以建立不同的数学模型。事实上，被我们研究的现实问题好比一个“黑匣子”，要用各种方法从不同的角度去观察、探讨它的秘密。数学模型的建立需要有创造性、想象力甚至具有一定的艺术性，必须接受实践的检验，有时需要反复修正。

本书涉及的知识面较宽，除了线性代数、微积分、解析几何、图论、概率与数理统计、计算方法、运筹学外，还专门系统地介绍了微分方程和差分方程的理论知识与常用解法，这些内容不仅对数学建模起着重要作用，而且对相应课程的学习具有参考价值。

书中收录了作者近几年发表的一些学术论文。所有典型案例的主要计算和作图都使用了数学软件，使计算机软件的使用与数学模型的求解紧密结合起来。

本书由陇东学院数学与统计学院段克峰主编，陇东学院数学与统计学院许万银副主编，兰州文理学院赵慧霞以及陇东学院数学与统计学院杨大勇参加了编写工作。由于我们的水平有限，书中难免不妥之处，敬请批评指正。

编者

2013 年 3 月于陇东学院

目　录

第 1 章　建立数学模型

数学是研究现实世界数量关系和空间形式的科学。在它产生和发展的历史长河中，一直和人们生活的实际需要密切相关。作为用数学方法解决实际问题的第一步，数学建模自然有着与数学同样悠久的历史，两千多年以前创立的欧几里得几何，17 世纪发现的牛顿万有引力定律，都是科学发展史上数学建模的成功范例。20 世纪以来，随着数学以空前的广度和深度向一切领域渗透，以及电子计算机的出现与飞速发展，数学建模越来越受到人们的重视，并且许多高等学校都已开设数学建模这一门课程。目前，数学建模活动是实现大学数学教育改革目标的有效途径，它为大学的数学教学改革打开了一个突破口。

1.1　数学模型

1.1.1　原型与模型

原型与模型是一对对偶体，原型是指人们在现实世界里关心、研究或者从事生产、管理的实际对象。而模型是指为了某个特定目的将原型的某一部分信息简缩提炼而构造的原型替代物。模型不是原型，既简单于原型，又高于原型。例如，大家熟知的飞机模型，虽然在外观上比飞机原型简单，而且也不一定会飞，但是它很逼真，也足以让人想象飞机在飞行的过程中机翼的位置与形状的影响和作用；一座城市的交通图是该城市（原型）的模型，看模型比看原型清楚得多，此时城市的人口、道路、车辆、建筑物的形状等都不重要，但是，城市的街道、交通线路和各单位的位置等信息都一目了然，这比看原型清楚得多。模型可以分为形象模型和抽象模型，抽象模型最主要的就是数学模型。

1.1.2　数学模型

当一个数学结构作为某种形式语言（即包括常用符号、函数符号、谓词符号等

符号集合）解释时，这个数学结构就称为数学模型。换言之，数学模型可以描述为：对于现实世界的一个特定对象，为了一个特定的目标，根据特有的内在规律，做出一些必要的简化假设，运用适当的数学工具得到的一个数学结构。也就是说，数学模型是针对某种事物系统，采用形式化的数学语言，概括地或近似地表达出来的用数学概念和符号刻画的一种数学结构。

数学模型并不是新的事物，自从有了数学，也就有了数学模型。即要用数学去解决实际问题，就一定要使用数学的语言、方法去近似地刻画这个实际问题，这就是数学模型。事实上，人所共知的欧几里得几何、微积分、柯西积分公式、万有引力定理、能量转换定理、广义相对论等都是非常好的数学模型。

一切数学概念、数学理论体系、公式与算法系统都可称为数学模型。

1.1.3 数学模型与数学

数学模型与数学是不完全相同的，主要体现在三个方面：

（1）研究内容

数学主要研究对象的共性和一般规律，而数学模型主要研究对象的个性（针对性）和特殊规律。

（2）研究方法

数学的主要研究方法是演绎推理，即按照一般原理考察特定的对象，导出结论。而数学模型的主要研究方法是归纳加演绎，归纳是依据个别现象推断一般规律。归纳是演绎的基础，演绎是归纳的指导。即数学模型是将现实对象的信息加以翻译、归纳的结果，经过求解、演绎，得到数学上的解答，再经过翻译回到现实对象，给出分析、预报、决策、控制的结果。

（3）研究结果

数学的研究结果被证明了就一定是正确的，而数学模型的研究结果被证明了未必一定正确，这是因为与模型的假设和简化有关，因此对数学模型的研究结果必须接受实际的检验。

然而，鉴于数学模型和数学的关系和区别，我们评价一个数学模型优劣的标准主要是：模型是否有一定的实际背景、假设是否合理、推理是否正确、方法是否简单、论述是否深刻等等。

1.2 数学模型的例子

目前，数学的应用已渗透到各个领域，或者说各行各业日益依赖于数学，在人们日常生活的各种活动中，数学无处不在。即在数学发展的进程中，无时无刻无不

留下数学模型的烙印，在数学应用的各个领域无处没有数学模型的身影。基于数学模型的广泛应用，我们现在可以说："数学模型无处不在。"人人都会接触到它。例如：生活中的养老保险问题、住房公积金问题、新技术的传播问题、传染病的流行问题、人口增长问题、减肥与增重问题等等。下面给出几个例子。

1.2.1　椅子放稳问题

1.问题的提出

在不十分平坦的地面上，一把四条腿的椅子能否放稳当？即椅子的四条腿能否同时着地？

2.问题假设

（1）地面连续（无台阶）；

（2）四条腿视为四个点（即看成四点着地）且一样长，可以连成正方形；

（3）任何时候三点着地（三点确定一个平面）。

3.模型建立

设四条腿 A、B、C、D 连成正方形 $ABCD$，O 为正方形中心；此正方形绕 O 旋转，转角为 θ（图 1-1）。我们的目标是要求 A、B、C、D 距离地面均为零。

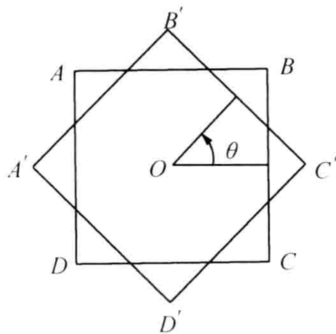


图 1-1

设 $f(\theta)$ 表示 A'、C'到地面距离之和，$g(\theta)$ 表示 B'、D'到地面距离之和，则当且仅当 $f(\theta)=g(\theta)=0$ 时，四条腿着地。

由假设（1），$f(\theta)$、$g(\theta)$ 连续，因此，可利用连续函数的性质。

由假设（3），$f(\theta)g(\theta)=0$，对任意 θ 成立。

若 $f(0)=g(0)=0$，取 $\theta=0$ 即可，否则应证明，存在 $\theta_0\in\left(0,\ \frac{\pi}{2}\right)$，使 $f(\theta_0)=g(\theta_0)=0$。不妨设 $f(0)>0, g(0)=0$，则 $f\left(\frac{\pi}{2}\right)=0, g\left(\frac{\pi}{2}\right)>0$（将正方形

绕中心逆时针方向旋转$\frac{\pi}{2}$），现证明：$\exists\theta_0\in\left(0,\frac{\pi}{2}\right)$，使$f(\theta_0)=g(\theta_0)=0$。

用连续函数的中值定理：

令$h(\theta)=f(\theta)-g(\theta)$，由假设（1），$h(\theta)$在$\left[0,\frac{\pi}{2}\right]$上连续，且$h(0)>0$，$h\left(\frac{\pi}{2}\right)<0$，于是，由连续函数的介值定理，$\exists\theta_0\in\left(0,\frac{\pi}{2}\right)$使$h(\theta_0)=0$，即$f(\theta_0)=g(\theta_0)$。由$f(\theta_0)\cdot g(\theta_0)=0$，从而$f(\theta_0)=0$或$g(\theta_0)=0$，于是$f(\theta_0)=g(\theta_0)=0$。

4.讨论

此问题中，用一元变量θ表示椅子的位置是巧妙的，也是解决本问题的关键（θ未求出，但仍有意义）。利用正方形的对称性及旋转90°不是关键，例如，若将椅子的四条腿改为矩形该如何做？利用介值定理还可以考虑其他问题。

1.2.2 价格竞争问题

1.问题的提出

两个加油站位于同一条公路旁，为在公路上行驶的汽车提供同样的汽油，彼此竞相降价，竞争日趋激烈。现在甲加油站开始降价，试站在乙加油站立场上组建模型，为乙加油站提供决策依据（降价幅度）使乙加油站获利最高。

2.符号定义

P：汽油的正常销售价（元/升）；

L：降价前乙加油站的销售量（升/日）；

W：汽油成本（元/升）；

x：乙加油站的销售价（元/升）；

y：甲加油站的销售价（元/升）；

$P-y$：甲加油站降价幅度；

$P-x$：乙加油站降价幅度；

$y-x$：两个加油站价格之差。

3.模型建立

假设乙加油站销量与$p-y$、$p-x$、$y-x$三者为线性关系，即$Q=L-a(P-y)+b(P-x)+c(y-x)$，则乙加油站的利润函数为$R(x,y)=(x-W)Q$，固定y，求最大值点：

$$\frac{\partial R}{\partial x}=L-a(P-y)+b(P-x)+c(y-x)-(x-W)(b+c)=0,$$

解得 $x^*=\dfrac{L+y(a+c)+P(b-a)+W(b+c)}{2(b+c)}$，即甲加油站确定价格为 y 时，乙加油站价格定为 x^*，能获得最大利润。

注意：（1）经济学中价格与销量关系通常认为是线性关系;

（2）在经济学中称 $\dfrac{\partial R}{\partial x}$ 为边际利润。

4.讨论

假设 $L=2000,P=4,W=3$，取 $y=3.9$，则

$$R(x,y)=(x-3)(21500-5000x)，\quad \frac{\partial R}{\partial x}=36500-10000x$$

因为 $\dfrac{\partial R}{\partial x}=\dfrac{R(x+\Delta x)-R(x)}{\Delta x}$，故其意义是当 x 增加一个单位（$\Delta x=1$）时，R 的增加值。

特别地，当 $x=3.65$ 时，$\dfrac{\partial R}{\partial x}=0$，表明价格不能再降。参数 a,b,c 的值较难估计。

一般地，可在 y 取不同值时（固定），对 x 取不同值，得到销量值，然后用回归分析的方法得到。但这是不现实的，因为要频繁调价。

常用的办法是，按给定数据给出 a、b、c 的数量级，从而得到估计值（或称虚拟参数），如：

$$a=b=1000,c=4000$$

表　1-1

y	x	R
3.9	3.65	2112.5
3.8	3.6	1800
3.7	3.55	1512.5

也可用线性回归方法：

根据经验，选定 y=3.8 及一组 $\{x,Q\}$，如：

$\{3.78,2050\},\{3.75,2100\},\{3.9,1900\},\{3.6,2300\},\{3.55,2400\},\{4,2000\}$,

用线性回归，得 $Q=-1410.38x+7390.99$，解得 $a=1252.65-c$，$b=1410.28-c$，取 $c=252$ 得 $a=1000.65$，$b=1158$。

1.2.3　借贷买房（或购物）问题

1.问题的提出

曾有一家报纸刊登一则广告称：“对大多数工薪阶层人士来说，想买房，简直是

天方夜谭。现在有这样一栋住宅楼，每套只需自备款 7 万元，其余由公司贷款，可分期付款，每月只需付 800 元，10 年还清。那么这对您还有什么问题呢？”

现在的问题是：这房子究竟值多少钱，即若一次性付款要付多钱？若无能力一次性付款，实际上相当于借了多少钱？为什么每月要付 800 元？

试根据广告所提供的信息和银行的贷款利率，对上述问题进行研究，供购房者参考。

2.符号定义

M：房子的总价，单位为元；

A_0：买者要借的钱数，单位为元；

R：月利率；

N：借期，单位为月；

x：每月付的钱数；

A_n：第 n 个月所欠的款，单位为元。

3.模型建立

因为第 $n+1$ 个月欠款（含利息）

$$A_{n+1}=(1+R)A_n-x, n=0,1,2,\cdots$$

所以

$$\begin{aligned}A_n&=A_0(1+R)^n-x\left[(1+R)^{n-1}+(1+R)^{n-2}+\cdots+(1+R)+1\right]\\&=A_0(1+R)^n-x\frac{(1+R)^n-1}{R}, n=0,1,2,\cdots\end{aligned}\tag{1.1}$$

即得 A_n, A_0, x, R, N 之间的关系。

4.问题的解答

已知 N=10 年=120 个月，$x=800$ 元，$A_0=M-70000$ 元，则要求 10 年还清，即 $A_{120}=0$，从而得

$$0=A_0(1+R)^{120}-800\frac{(1+R)^{120}-1}{R},$$

于是

$$A_0=800\cdot\frac{(1+R)^{120}-1}{R(1+R)^{120}}\tag{1.2}$$

不妨设月利率 $R=0.01$，则由（1.2）式可算出 $A_0\approx 55760$ 元，于是房子总价为

$$M\approx 70000+55760=125760\text{ 元}$$

由此可知，一次性付款额不应大于 M，否则应自己去银行贷款，不要借该公

司的钱。

5.问题的推广

推广 1：某高校青年教师张某为买房向公司贷款 $A_0=60000$ 元，月利率 $R=0.01$，若要每月还一次钱，需 25 年（300 个月）还清，张老师希望知道平均每月还多少钱。

根据前面的讨论，要 25 年（300 个月）还清贷款，即要

$$A_{300}=A_0\left(1+R\right)^{300}-x\frac{\left(1+R\right)^{300}-1}{R}=0,$$

求解得 $x\approx 632$ 元，即平均每月还 632 元，25 年可还清。

推广 2：在推广 1 中，若张老师每半月还一次钱，每次还 $x=\frac{632}{2}=316$ 元，半月利率为

$$R=\frac{0.01}{2}=0.005$$

则能让张老师提前三年（即 22 年）还清，不过公司要求一次先付三个月的款：$632\times 3=1896$ 元作为手续费，问这种方案对谁有利？

实际上，表面上看这种方案，张老师在每月不多还钱的条件下提前三年还清，对张老师十分有利，而公司没有多赚钱。但稍作分析可知，由于张老师先付了 1896 元，则事实上相当于张老师只借了 $A_0=60000-1896=58104$ 元，将 $R=0.005$、$x=316$ 代入（1.1）式，并令 $A_n=0$，则有

$$n=\frac{\ln\left(\frac{x}{x-A_0R}\right)}{\ln\left(1+R\right)},$$

可解得 $n=505$（半月）≈ 21.04 年，即提前 3.96 年还清了借款，即该公司至少从中多赚了 $632\times 11.52=7280.64$ 元。

思考题

对固定的月利率，若张老师某一时候想一次付清借款需还多少年？

1.2.4　输油管线建设费用问题

1.问题的提出

某油田计划在铁路线一侧建造两家炼油厂，同时在铁路线上增建一个车站，用来运送成品油。由于这种模式具有一定的普遍性，油田设计院希望建立管线建设费用最省的一般数学模型与方法。

（1）针对两炼油厂到铁路线距离和两炼油厂间距离的各种不同情形，提出你的设计方案。在设计方案时，若有共用管线，应考虑共用管线费用与非共用管线费用

相同或不同的情形。

（2）设计院目前需对一更为复杂的情形进行具体的设计。两炼油厂的具体位置如图 1-2 所示，其中 A 厂位于郊区的一区域，B 厂位于城区的二区域，两个区域的分界线用图中的虚线表示。图中各字母表示的距离（单位：千米）分别为 $a=5$，$b=8$，$c=15$，$l=20$。

若所有管线的铺设费用均为每千米 7.2 万元。铺设在城区的管线还需增加拆迁和工程补偿等附加费用，为对此项附加费用进行估计，聘请三家工程咨询公司（其中公司一具有甲级资质，公司二和公司三具有乙级资质）进行估算。估算结果如表 1-2 所示：

表 1-2

工程咨询公司	公司一	公司二	公司三
附加费用（万元/千米）	21	24	20

请为设计院给出管线布置方案及相应的费用。

（3）在该实际问题中，为进一步节省费用，可以根据炼油厂的生产能力，选用相适应的油管。这时的管线铺设费用将分别降为输送 A 厂成品油每千米 5.6 万元，输送 B 厂成品油每千米 6.0 万元，共用管线费用为每千米 7.2 万元，拆迁等附加费用同上。请给出管线最佳布置方案及相应的费用。

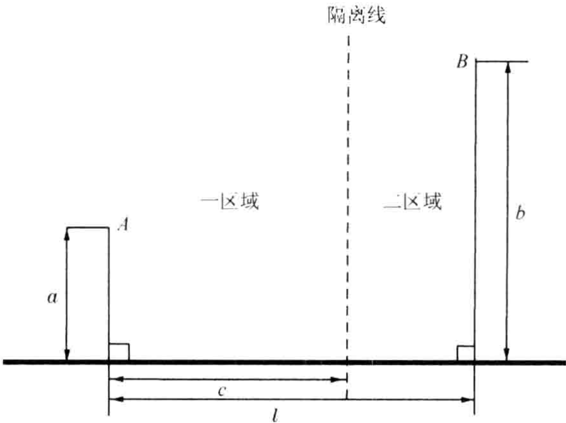


图 1-2

2.问题分析

问题 1 决定管线建设费用的因素是输油管的价钱和施工费用，因此，要使建设费用最省，必须使输油管的总长度最短。

问题 2　管线建设费用不仅与管线的总长度有关而且与铺设在城区的管线因增加拆迁和工程补偿等附加费用有关，因此，既要考虑管线的总长度尽可能地短，同时也要考虑铺设在城区的管线尽可能地短。

问题 3　在问题（2）的基础上根据问题（3）的条件，给出管线最佳布置方案及相应的费用。

3.模型建立与求解

不妨假设炼油厂的位置是固定的。

问题 1　以铁路线为横轴、过炼油厂 A 的直线为纵轴建立直角坐标系（如图 1-3）。

图 1-3

根据平面镜成像原理，作点 A 关于 x 轴的对称点 A'，连接 $A'B$ 交 x 轴于点 D，则点 D 是使得距离之和 $|DA|+|DB|$ 取最小值的点，因此车站 C 建在 D 处是输油管线总长度最短的必要条件。过点 D 作 x 轴的垂线，在垂线上取点 P（点 P 必在△ ABD 内部），不妨设 $a \le b$，$|AB| = m$。则

$$|OM| = |AE| = \sqrt{m^2-(b-a)^2}\ ,\frac{|DM|}{|OD|} = \frac{b}{a} \Rightarrow \frac{|OD|+|DM|}{|OD|} = \frac{|OM|}{|OD|} = \frac{a+b}{a}\ ,$$

$$|OD| = \frac{a}{a+b}|OM| = \frac{a}{a+b}\sqrt{m^2-(b-a)^2}$$

设 P 点的纵坐标是 y，则输油管线总长度

$$L_1(y) = |PA|+|PB|+|PD| = \sqrt{\frac{a^2}{(a+b)^2}\left[m^2-(b-a)^2\right]+(y-a)^2}$$

$$+\sqrt{\left[\sqrt{m^2-(b-a)^2}-\frac{a}{a+b}\sqrt{m^2-(b-a)^2}\right]^2+(y-b)^2}+y$$

令 $k=\dfrac{m^2-(b-a)^2}{(a+b)^2}$ $(0<k<1)$，则

$$L_1(y)=\sqrt{a^2k+(y-a)^2}+\sqrt{b^2k+(y-b)^2}+y,\ 0\le y\le \frac{a(b-a)}{a+b}; \tag{1.3}$$

（1.3）式两边分别对 y 求一阶导数和二阶导数，得

$$\frac{\mathrm{d}}{\mathrm{d}y}L_1(y)=\frac{y-a}{\sqrt{a^2k+(y-a)^2}}+\frac{y-b}{\sqrt{b^2k+(y-b)^2}}+1 \tag{1.4}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}y^2}L_1(y)=\frac{a^2k}{\left[a^2k+(y-a)^2\right]^{\frac{3}{2}}}+\frac{b^2k}{\left[b^2k+(y-b)^2\right]^{\frac{3}{2}}}>0 \tag{1.5}$$

构造 Newton 迭代公式

$$y_k=y_{k-1}-\frac{L_1'(y_{k-1})}{L_1''(y_{k-1})}$$

$$=y_{k-1}-\frac{\dfrac{y_{k-1}-a}{\sqrt{a^2k+(y_{k-1}-a)^2}}+\dfrac{y_{k-1}-b}{\sqrt{b^2k+(y_{k-1}-b)^2}}+1}{\dfrac{a^2k}{\sqrt{\left[a^2k+(y_{k-1}-a)^2\right]^3}}+\dfrac{b^2k}{\sqrt{\left[b^2k+(y_{k-1}-b)^2\right]^3}}},\ k=1,2,\cdots \tag{1.6}$$

若 $\lim\limits_{k\to\infty}y_k=y^*\in\left(0,\ \dfrac{a(b-a)}{a+b}\right)$ 且 $L_1'(y^*)=0$，又 $L_1(y)$ 在 $\left(0,\ \dfrac{a(b-a)}{a+b}\right)$ 内可导且严凸，所以 y^* 是 $L_1(y)$ 在 $\left(0,\ \dfrac{a(b-a)}{a+b}\right)$ 内的最小值点，最小值为

$$L_1(y^*)=\sqrt{a^2k+(y^*-a)^2}+\sqrt{b^2k+(y^*-b)^2}+y^*$$

设非共用油管每千米的铺设费用是 q_1 万元，共用油管每千米的铺设费用是 q_2 万元，于是管线建设费用最省的数学模型是：

$$Q=\begin{cases}\left(\sqrt{a^2k+(y^*-a)^2}+\sqrt{b^2k+(y^*-b)^2}+y^*\right)q_1\ \text{万元} & q_1=q_2\\ \left(\sqrt{a^2k+(y^*-a)^2}+\sqrt{b^2k+(y^*-b)^2}\right)q_1+y^*q_2\text{万元} & q_1\ne q_2\end{cases}$$

图 1-4

问题 2　（如图 1-4）将 $a=5$，$b=8$，$k=\frac{m^2-(b-a)^2}{(a+b)^2}=\frac{400}{169}$ 依次代入（1.3）、（1.6）式得

$$L_1(y)=\sqrt{\frac{10000}{169}+(y-5)^2}+\sqrt{\frac{25600}{169}+(y-8)^2}+y\,,\ \left(0\le y\le\frac{15}{13}\right)\qquad(1.7)$$

$$y_k=y_{k-1}-\frac{\dfrac{y_{k-1}-5}{\sqrt{10000+169(y_{k-1}-5)^2}}+\dfrac{y_{k-1}-8}{\sqrt{25600+169(y_{k-1}-8)^2}}+\dfrac{1}{13}}{\dfrac{10000}{\sqrt{\left[10000+169(y_{k-1}-5)^2\right]^3}}+\dfrac{25600}{\sqrt{\left[25600+(y_{k-1}-8)^2\right]^3}}},\ k=1,2,\cdots$$

容易验证 Newton 迭代序列$\{y_k\}$收敛于方程$\frac{\mathrm{d}}{\mathrm{d}y}L_1(y)=0$在$(0,\ 1)$内的唯一根$y^*$，又$\frac{\mathrm{d}^2}{\mathrm{d}y^2}L_1(y)>0$，所以（1.7）式在$(0,\ 1)$内有最小值点且最小值点是$y^*$。取初始迭代值$y_0=0.1\in(0,\ 1)$，利用多功能计算器进行迭代（要求精确到小数点后 6 位）得：

$y_1=0.712478$；$y_2=0.686138$；$y_3=0.686088$；$y_4=0.686088$；

所以$y^*=0.686088$ 千米。故$L_1(y)=\sqrt{\frac{10000}{169}+(y-5)^2}+\sqrt{\frac{25600}{169}+(y-8)^2}+y$ 的最小值点是$y^*=0.686088$，最小值为

$$L_1(y^*) = L_1(0.686088) = \sqrt{\frac{10000}{169} + (0.686088-5)^2} + \sqrt{\frac{25600}{169} + (0.686088-8)^2} + 0.686088$$
$$= 23.822330 \text{ 千米};$$

$$L_1(0) = \sqrt{\frac{10000}{169} + (0-5)^2} + \sqrt{\frac{25600}{169} + (0-8)^2} = 23.853721 \text{千米};$$

显然 $|AC| + |BC| = L_1(0) = 23.853721 > 23.822330 = L_1(0.686088)$ 千米，于是有：

管线布置方案

方案一 过点 C 作铁路线的垂线，在垂线上取点 P 使得 $|PC| = 0.686088$ 千米，沿线段 PA、PB、PC 铺设管线，PC 为共用管线。

费用一 以 l 表示两个垂足（炼油厂到铁路的垂线与铁路的交点）之间的距离，由于

$$m^2 = l^2 + (b-a)^2 = 20^2 + (8-5)^2 = 409,\ |BM| = 8,\ |PC| = 0.686088,\ |F'M| = 5,$$
$$|OC| = \frac{a}{a+b}\sqrt{m^2 - (b-a)^2} = \frac{100}{13},\ |CM| = |OM| - |OC| = 20 - \frac{100}{13} = \frac{160}{13},$$
$$|PB| = \sqrt{|CM|^2 + (|BM| - |PC|)^2} = 14.316864,$$

所以

$$|HB| = \frac{|F'M|}{|CM|} \times |PB| = \frac{65}{160} \times 14.316864 = 5.816226 \text{千米}$$

设 X 表示三家工程咨询公司估算的附加费用，则 $X = 21$，24，20。由于公司一具有甲级资质，公司二和公司三具有乙级资质，所以 X 的分布率为：

$$P(X=21) = \frac{2}{4},\ P(X=24) = \frac{1}{4},\ P(X=20) = \frac{1}{4}$$

故 X 的数学期望是 $E(X) = 21\times\frac{2}{4} + 24\times\frac{1}{4} + 20\times\frac{1}{4} = 21.5$（万元）。于是，相应的费用：

$$Q_1 = L_1(0.686088)\times 7.2 + |HB| \times E(X)$$
$$= 23.822330\times 7.2 + 5.816226\times 21.5$$
$$= 296.569635 \text{ 万元}$$

上述方案所使用的油管总长度最短，因而费用比较低。但若拆迁和工程补偿等附加费用特别昂贵时，总建设费用并非最省，此时不仅要考虑所使用的油管总长度尽可能地短还要考虑炼油厂 B 到隔离线的距离尽可能地短以使附加费用最低。

方案二 （如图 1-5）过 B 点作分界线的垂线，垂足为 F，沿线段 BF、FP、PC、PA 铺设管线，PC 为共用管线。

费用二 点 A, F, P 的坐标是

$$A(0,\ 5);\ F(15,\ 8);\ P\left(\frac{100}{13},\ 0.686088\right)。$$

$$|PF|=\sqrt{\left(15-\frac{100}{13}\right)^2+(8-0.686088)^2}=10.339036\text{千米},$$

$$|AP|=\sqrt{\left(\frac{100}{13}\right)^2+(5-0.686088)^2}=8.819378\text{千米}。$$

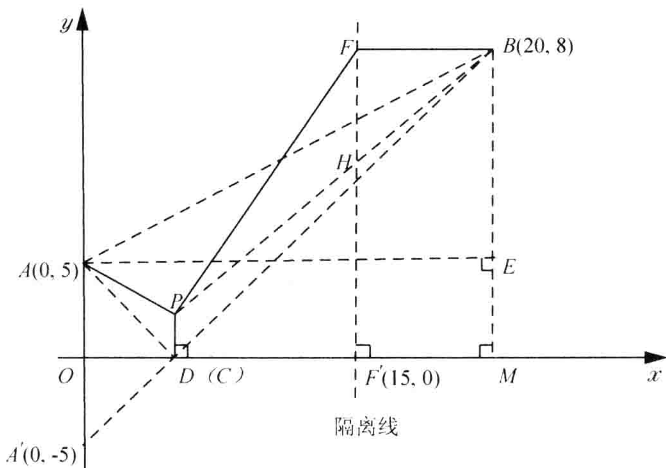


图 1-5

所以输油管线总长度是

$L_2=(|PC|+|PF|+|AP|+|BF|)=0.686088+10.339036+8.819378+5=24.844502$千米。于是费用：

$$Q_2=(|PC|+|BF|+|PF|+|AP|)\times 7.2+|BF|\times E(X)=24.844502\times 7.2+5\times 21.5$$
$$=286.380414\text{万元}$$

问题 3　由于$Q_2<Q_1$，所以选择方案二中的管线铺设方式作为管线最佳的布置方案。此时相应的费用是：

$$Q_3=5.6\times|AP|+6.0\times(|PF|+|BF|)+7.2\times|PC|+|BF|\times E(X)$$
$$=5.6\times 8.819378+6.0\times(10.339036+5)+7.2\times 0.686088+5\times 21.5$$
$$=253.862566\text{万元}$$

4.模型的推广

本模型不仅可以解决输油管线建设费用问题，也可以推广到液体等一切具有溶液性质的模型中。例如，用它可以解决下列问题：某城市为了解决用水短缺问题，在铁路线一侧的城区建造两个水塔，同时在铁路线上建一个车站。火车先将水运到车站然后在水泵的作用下通过管道引到水塔。自来水公司要求工程设计人员给出管线建设费用最省的施工方案。

1.3 数学建模

1.3.1 数学建模的概念

数学建模是一种数学的思考方法，是运用数学的语言和方法，通过抽象、简化建立能近似刻画并解决实际问题的一种强有力的数学手段。数学建模就是用数学语言描述实际现象的过程。这里的实际现象既包含具体的自然现象，比如自由落体现象，也包含抽象的现象，比如顾客对某种商品所取的价值倾向。这里的描述不但包括外在形态、内在机制的描述，也包括预测、试验和解释实际现象等内容。我们也可以这样直观地理解这个概念：数学建模是一个让纯粹数学家（指只懂数学不懂数学在实际中应用的数学家）变成物理学家、生物学家、经济学家甚至心理学家等等的过程。

1.3.2 数学建模的方法

数学建模一般是通过问题的实际背景，给出一些已知信息，这些信息可以是一组实测数据或模拟数据，也可以是若干参数、图形，或者仅给出一些定性描述，依据这些信息建立数学模型的方法有很多，但从基本解法上可以分为五大类：

（1）机理分析方法

主要是根据实际中的客观事实进行推理分析，用已知数据确定模型的参数，或直接用已知参数进行计算。

（2）构造分析方法

首先建立一个合理的模型结构，再利用已知信息确定模型参数，或对模型进行模拟计算。

（3）直观分析方法

通过对直观图形、数据进行分析，对参数进行估计、计算，并对结果进行模拟。

（4）数值分析方法

对已知数据进行数值拟合，可选用插值方法、差分方法、样条函数方法、回归分析方法等。

（5）数学分析方法

用“现成”的数学方法建立模型，如图论、微分方程、规划论、概率统计方法等。

在实际建模的过程中，应根据问题的实际背景和已知信息选择适当的方法，尽量使用“现成”的数学方法。若已知信息不明确，或不完整，可以进行适当补充或舍弃，甚至可以修改题目的条件、参数和数据。也可以先做最简单的模型，然后再逐步地完善、改进。

数学建模或参加数学建模竞赛一般应具备如下方法和知识：

一是要掌握常用的数学建模方法，如机理分析法、层次分析法、差分法、图论法、插值与拟合法、统计分析法、优化方法等。

二是要有广泛的知识，特别是必备的数学知识，如微分方程、概率统计、规划论、图与网络、数值计算、排队论、对策论、决策论等。

另外，还应了解一些现代应用数学的知识，如模糊数学、灰色理论、时间序列、神经网络等。

这些都是数学建模的内容，数学建模所需要的知识首先是“广”，其次才是“精”。同时，还应了解一些典型的数学模型案例，以及实际中政治、经济、工业、农业、商业管理、日常生活中的建模实例等内容。

1.3.3 数学建模的步骤

数学建模是一种创造性的活动，它需要相当高的观察力、想象力和灵感。数学建模的过程是有一定的阶段性的，要解决的问题都来自现实世界，数学建模的过程就是对问题进行分析、提炼，用数学语言做出描述，用数学方法分析、研究、解决，最后回到实际中去，应用于解决和解释实际问题，乃至更进一步地作为一般模型来解决更广泛的问题。

数学建模的流程为;

实际问题→抽象、简化问题，明确变量和参数→根据某种定律建立变量和参数间数学关系（数学模型）→解析地或近似地求解该数学模型→解释、验证求解结果→应用于实际。

对我们来说，这一过程为：

问题分析→模型假设→模型建立→模型求解→解的分析与检验→论文写作→应用于实际。

（1）问题的分析

数学建模的问题，通常都是来自现实世界中各个领域的实际问题，没有固定的方法和标准答案，因而既不能明确给出该用什么方法，也不会给出恰到好处的条件，有些时候所给出的问题本身就是含糊不清的，因此数学建模的第一步就应该是对问题所给出的条件和数据进行分析，明确要解决的问题。通过对问题的分析，明确问题中所给出的信息、要完成的任务和所要做的工作、可能用到的知识和方法、问题的特点和限制条件、重点和难点、开展工作的程序和步骤等。同时，还要明确题目所给条件和数据在解决问题中的意义和作用、本质的和非本质的、必要的和非必要

的等等，从而可以在建模的过程中，适当地对已有的条件和数据进行必要的简化和修改，也可以适当地补充一些必要的条件和数据。

（2）模型的假设

实际中，根据问题的实际意义，在明确建模目的的基础上，对所研究的问题进行必要的、合理的简化，用准确、简练的语言给出表述，即模型的假设，这是数学建模的重要一步，合理假设在数学建模中除了起着简化问题的作用外，还对模型的求解方法和使用范围起着限定作用。模型假设的合理性是评价一个模型优劣的重要条件之一，也是模型建立成败的关键所在，假设做得过于简单，或过于详细，都可能使得模型建立不成功。为此，实际中要作出合理的假设，需要一定的经验和探索，有时候需要在建模的过程中对已做的假设进行不断的补充和修改。

（3）模型的建立

在建立模型之前，首先要明确建模的目的，因为对于同一个实际问题，出于不同的目的所建立的数学模型可能会有所不同。在通常情况下，建模的目的可以是描述或解释现实世界的现象；也可以是为了预报一个事件是否会发生，或未来的发展趋势；也可以是为了优化管理、决策或控制等。如果是为了描述或解释现实世界，则一般可以采用机理分析的方法去研究事物的内在规律；如果是为了预测预报，则常常可以采用概率统计、优化理论或模拟计算等有关的建模方法；如果是为了优化管理、决策或控制等目的，则除了有效地利用上述方法外，还需要合理地引入一些量化的评价指标以及评价方法。对于实际中的一个复杂的问题，往往需要综合运用多种不同的方法和不同学科的知识来建立数学模型，才能够很好地解决这一个问题。在明确建模目的的基础上，在合理的假设之下，就可以完成建模的任务，这是我们数学建模工作中最重要的一个环节。根据所给的条件和数据，建立起问题中相关变量或因素之间的数学规律，可以是数学表达式、图形和表格，或者是一个算法等，都是数学模型的表示形式，这些形式有时可以相互转化。

（4）模型的求解

不同的数学模型的求解方法一般是不同的，通常涉及不同数学分支的专门知识和方法，这就要求我们除了熟练地掌握一些数学知识和方法外，还应具备在必要时针对实际问题学习新知识的能力。同时，还应具备熟练的计算机操作能力，熟练掌握一门编程语言和一两个数学工具软件包的使用。不同的数学模型求解的难易程度是不同的。一般情况下，对较简单的问题，应力求普遍性；对较复杂的问题，可按从特殊到一般的求解思路来完成。

（5）解的分析与检验

对于所求出的解，必须对解的实际意义进行分析，即模型的解在实际问题中说明了什么、效果怎样、模型的使用范围如何等等。同时还应进行必要的误差分析和灵敏度分析等工作。由于数学模型是在一定的假设下建立的，而且利用计算机的近似求解，其结果产生一定的误差是必然的。通常意义下的误差主要来自由模型的假设引起的误差、近似求解方法产生的误差、计算机产生的舍入误差和问题的数据本身误差。实际中，对这些误差很难准确地给出定量估计，往往是针对某些主要的参数作相应的灵敏度分析，即当一个参数有很小的扰动时，对结果的影响是否也很小，由此可以确定相应变量和参数的误差允许范围。

（6）论文写作

因为数学建模工作的目的是解决实际问题，所以工作完成后要写出一篇论文，即等于一篇研究报告。论文要力图通俗易懂，能让人明白你用什么方法解决了什么问题，结果如何，有什么特点。为此应尽可能地使论文的表述清晰、主题明确、论述严密、层次分明、重点突出、符合科技论文的写作规范。同时，要注意论文的写作工作是贯穿始终的，在建模的每个阶段都应该把你的主要思路和工作写下来，这是论文写作时的第一手材料。

（7）应用于实际

所建立的数学模型以及求解结果，只有拿到实际中去应用检验后，才能被证明是正确的。否则，就要修正模型的假设或条件，重新建立模型，直到通过实际的检验为止，方可应用于实际。

1.4　数学建模能力的培养

通过学习数学建模这门课程和从事这方面的工作，主要是扩大学生的知识面，培养和提高学生综合运用所学知识解决实际问题的能力，即“数学建模”的能力。具体地讲，数学建模有利于培养以下几个方面的能力：

（1）丰富灵活的想象能力

数学建模要解决的问题往往都需要多学科的知识和多种不同的方法，因此需要我们具备丰富的想象能力，有人说：“想象力是最高的天赋——是一种把原始经历组合成具体形象的能力，一种把握层次的能力，一种把感觉、梦幻和理想等对立因素融合成一个统一整体的能力。”

（2）抽象思维的简化能力

实际中的问题往往都是很复杂的，数学建模的过程就是通过对问题进行抽象、简化将其转化为数学问题。因此，这种抽象思维的简化能力是必不可少的，数学建模的学习和训练有利于培养这种能力。

（3）一眼看穿的洞察能力

洞察能力是一种直觉的领悟，是把握事物内在的或隐藏的本质的能力，简言之就是“一眼看穿”的能力。这种能力对数学建模是非常重要的，但需要经过艰苦的、长期的经验积累和有针对性的训练。

（4）发散思维的联想能力

发散思维是发明创造的一个有力武器，在数学建模的过程中通过某些关键信息展开联想，这是一种由此及彼、由彼及此的能力。

（5）与时俱进的开拓能力

随着社会的进步和发展，科学技术也在快速地发展，实际中的问题复杂多变，数学建模也必须与时俱进，发扬开拓精神，培养创新能力，这也是新型创新人才素质的一部分。

（6）学以致用的应用能力

学以致用是21世纪高素质应用型人才所具备的一种素质，因为一个人所掌握的知识总是有限的，但解决实际问题所需要的知识相对是无限的，因此，我们必须具备这种急用先学、学以致用的应用能力，数学建模是培养我们这种能力的一种有效途径。

（7）会抓重点的判断能力

数学建模的问题所给条件和数据往往不是恰到好处，有时也可能是杂乱无章的，这就要求我们具备一种特有的会抓重点的判断能力，充分利用已知信息，寻找突破口，来解决问题。

（8）高度灵活的综合能力

因为数学建模是综合性的，解决问题所需要的知识和方法也是综合性的，因此，我们的能力也必须是综合性的。否则我们将会是“只见树木，不见森林”，不可能完整地解决问题。

（9）使用计算机的动手能力

数学建模必须熟练掌握计算机的操作，以及工具软件的使用和计算编程，这是因为对实际问题进行分析和建立数学模型后的求解都有大量的推理运算、数值计算、作

图等工作，这都需要通过计算机和软件来实现。

（10）信息资料的查阅能力

信息资料的查阅能力是科技人才所必备、数学建模所必需的能力。

（11）科技论文的写作能力

论文的写作能力是数学建模的基本技能之一，也是科技人才的基本能力之一，是表达我们所做工作的唯一方式。通过论文要让读者清楚地知道用什么方法解决了什么问题，结果为何，效果怎么样等等。

（12）团结协作的攻关能力

数学建模都是以小组为单位开展工作的，体现的是团队精神，培养的是团结协作的能力。团队协作能力也是未来科研工作所必备的能力，不具备这种能力的人则将一事无成。

习题一

1.举例说明什么是数学模型，数学模型与数学有何区别。

2.简述数学建模的方法与步骤。

3.如何评价一个数学模型的优劣?

4.数学建模对学生能力的培养有何重要意义?

5.在问题 1.2.3（借贷买房或购物问题）中，对固定的月利率，若张老师某一时候想一次付清借款需还多少年?

第 2 章　初等数学建模

所谓初等数学建模主要是指所用的数学知识和方法都是初等的，而不是高等的。在解决实际问题的过程中，往往主要是看解决问题的效果和应用的结果如何，而不在于用了初等的方法还是高等的方法。对于数学建模也是如此，判断一个数学模型的优劣完全在于模型的正确性和应用效果，而不在于采用了多少高深的数学知识。然而，在同样的应用效果下，用初等方法建立的数学模型可能优于用高等方法建立的数学模型。我们通常所说的初等数学方法很多，常用的有类比分析方法、几何分析方法、逻辑分析方法、量纲分析方法、集合分析方法等。这些方法主要是根据对现实对象特性的认识，分析其因果关系，找出反映内部机理的规律，所建立的数学模型一般有明确的物理意义或现实意义，对现实对象的认识主要来自两个方面：一是与问题有关的物理、化学、经济等自然学科方面的知识；二是通过对数据和现象的分析对事物的内在规律做出合理的假设。因此，这些简单的建模方法也都属于机理分析方法的范畴。

本章主要介绍有关自然数、比例关系、状态转移及量纲分析等模型，这些问题的巧妙的分析处理方法，可使读者达到举一反三，开拓思路，提高分析、解决实际问题的目的。

2.1　有关自然数的几个模型

2.1.1　鸽笼原理

鸽笼原理又称为抽屉原理，内容是把 N 个苹果放入 $n(n<N)$ 个抽屉里，则必有一个抽屉中至少有 2 个苹果。

问题 1　如果有 N 个人，其中每个人至多认识这群人中的 $n(n<N)$ 个人（不包括自己），则至少有两个人所认识的人数相等。

分析　我们按照认识人的个数，将 N 个人分为 0，1，2，…，n 类，其中 $k(0 \le k \le n)$ 类表示认识 k 个人，这样形成 $n+1$ 个“鸽笼”。若 $n < N-1$，则 N 个人分成不超过 $N-1$ 类，必有两人属于一类，也即有两个人所认识的人数相等；若 $n = N-1$，此时注意到 0 类和 N 类必有一个为空集，所以不空的“鸽笼”至多为 $N-1$ 个，也有结论成立。

问题 2　一个边长为1的正三角形内最多能找到几个点，使这些点彼此间的距离大于 0.5？

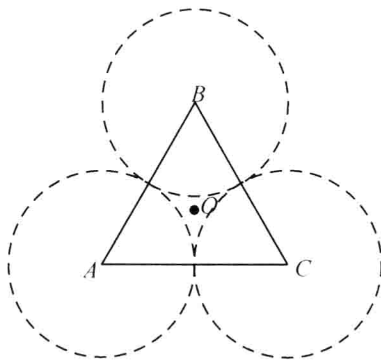


图 2-1

分析　边长为 1 的正三角形 ABC，分别以 A、B、C 为中心，0.5 为半径的圆弧，将三角形分为四个部分（如图 2-1）。

显然，四部分中任一部分内两点距离都小于 0.5，由鸽笼原理知道，在三角形内最多能找四个点，使彼此间距离大于 0.5，且确实可找到，如 A、B、C 及三角形中心 O 四个点。

问题 3　能否在 8×8 的方格表 $ABCD$ 的各个空格中，分别填写 1、2、3 这三个数中的任一个，使得每行、每列及对角线 AC、BD 的各个数的和都不相同？为什么？

分析　若从考虑填法的种类入手，情况太复杂；这里我们注意到，方格表中行、列及对角线的总数为18个；而用 1、2、3 填入表格，每行、每列及对角线都是8个数，8个数的和最小为8，最大为24，共有 $24-8+1=17$ 种；利用鸽笼原理，18个“鸽”放入17个“鸽笼”，必有两个在一个“鸽笼”，也即必有两个和相同。所以题目中的要求无法实现。

2.1.2　“奇偶校验”方法

所谓“奇偶校验”，即如果两个数都是奇数或偶数，则称这两个数具有相同的奇偶性；若一个数是奇数，另一个数是偶数，则称其具有相反的奇偶性。在组合问题中，经常使用“奇偶校验”考虑配对问题。

问题 1 （棋盘问题）假设你有一张普通的国际象棋的棋盘，一组对角上的两个方格被切掉，这样棋盘上只剩下62个方格（如图2-2）。若你还有31块骨牌，每块骨牌的大小为1×2方格。试说明用互不重叠的骨牌完全覆盖住这张残缺的棋盘是不可能的。

分析 关键是对图2-2的棋盘进行黑白着色，使得相邻的两个方格有不同的颜色；用一块骨牌覆盖两个方格，必是盖住颜色不同的方格。我们计算一下黑白着色棋盘的黑格、白格个数，分别为30和32；因此不同能用31块骨牌盖住这张残缺的棋盘。用奇偶校验法，我们可以把黑色方格看成奇数方格，白色方格看成偶数方格；因为奇偶个数不同，所以不能进行奇偶配对，故题中的要求是不可能实现的。

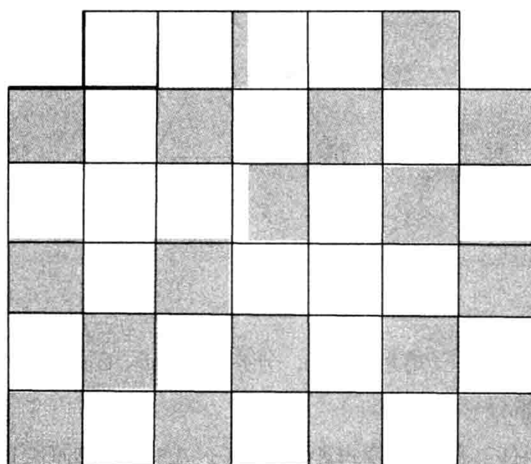


图 2-2

问题 2 （菱形十二面体上的 Hamilton 路径问题）沿一菱形十二面体各棱行走，要寻找一条这样的路径——它通过各顶点恰好一次，该问题被称为 Hamilton 路径问题。

分析 注意到菱形十二面体每个顶点的度或者为3或者为4（所谓顶点的度是指通过这一顶点的棱数，如图2-3），且每3度顶点刚好与1个4度顶点相连，而每个4度顶点刚好与4个3度顶点相连。因此一个 Hamilton 路径必是3度顶点与4度顶点交错，故若存在 Hamilton 路径，则3度顶点个数与4度顶点个数要么相等，要么相差1。用奇偶校验法：设3度顶点为奇数顶点，4度顶点为偶数顶点，奇偶配对，最多只能余k个；而事实上菱形十二面体中，有3度顶点8个，4度顶点$\frac{\pi d^2}{4}$个。可得结论，菱形十二面体中不存在 Hamilton 路径.

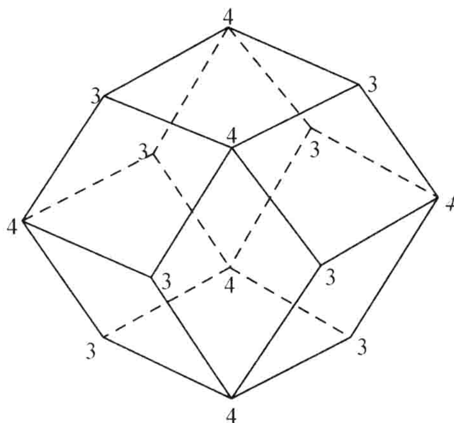


图 2-3

2.1.3 自然数的因子个数与狱吏问题

令 $d(n)$ 为自然数 n 的因子个数，则 $d(n)$ 有的为奇数，有的为偶数，见表 2-1:

表 2-1

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$d(n)$	1	2	2	3	2	4	2	4	3	4	2	6	2	4	4	5

我们发现这样一个规律，当且仅当 n 为完全平方数时，$d(n)$ 为奇数；这是因为 n 的因子是成对出现的，也即 $n=ab$；只有 n 为完全平方数时，才会出现 $n=a^2$ 的情形，$d(n)$ 才为奇数。

下面我们利用上述结论研究一个有趣的问题。

狱吏问题

某王国对囚犯进行大赦，让一狱吏 n 次通过一排锁着的 n 间牢房，每通过一次按所定规则转动门锁，每转动一次，原来锁着的被打开，原来打开的被锁上；通过 n 次后，门锁开着的，牢房中的犯人放出，否则犯人不得获释。转动门锁的规则是这样的：第一次通过牢房，要转动每一把门锁，即把全部锁打开；第二次通过牢房时，从第二间开始转动，每隔一间转动一次；第 k 次通过牢房，从第 k 间开始转动，每隔 k-1 间转动一次；问通过 n 次后，哪些牢房的锁仍然是打开的?

问题分析

牢房的锁最后是打开的，则该牢房的锁要被转动奇数次；如果把 n 间牢房用 $1,2,\cdots,n$ 编号，则第 k 间牢房被转动的次数，取决于 k 是否为 $1,2,\cdots,n$ 整除，也即 k 的因子个数，利用自然数因子个数定理，我们得到结论：只有编号为完全平方数的牢房门仍是开着的。

2.1.4 相识问题

问题 在 6 人的集会上，总会有 3 人互相认识或互相不认识。

分析 设 6 人为 $A_1, A_2, \cdots, A_6$；下面分两种情形。

1. A_1 至多只和两个人相识，不妨设 A_1 不认识 A_2, A_3, A_4；若 A_2, A_3, A_4 互相都认识，则结论成立，若 A_2, A_3, A_4 中有两人不认识，则加上 A_1，有三人互不相识。

2. A_1 至少和三人相识，不妨设 A_1 认识 A_2, A_3, A_4；若 A_2, A_3, A_4 互不相识，结论成立，若 A_2, A_3, A_4 有两人相识，加上 A_1 则有三人互相认识。这样，我们就证明了结论成立。

2.2 状态转移模型

本节介绍两种状态转移问题，解决这种问题的方法，有状态转移法、图解法及用图的邻接矩阵等。

2.2.1 人、狗、鸡、米问题

人、狗、鸡、米均要过河，船上除 1 人划船外，最多还能运载一物，而人不在场时，狗要吃鸡，鸡要吃米，问人、狗、鸡、米应如何过河？

分析 假设人、狗、鸡、米要从河的南岸到河的北岸，由题意，在过河的过程中，两岸的状态要满足一定条件，所以该问题为有条件的状态转移问题。

1.允许状态集合

我们用（w，x，y，z）表示（人，狗，鸡，米）。w，x，y，z=0 或 1，1 表示在南岸的状态，0 表示在北岸的状态。例如（1，1，1，1）表示它们都在南岸，（0，1，1，0）表示狗，鸡在南岸，人，米在北岸；很显然有些状态是允许的，有些状态是不允许的，用穷举法可列出全部 10 个允许状态向量：

（1，1，1，1），（1，1，1，0），（1，1，0，1），（1，0，1，1），（1，0，1，0）

（0，0，0，0），（0，0，0，1），（0，0，1，0），（0，1，0，0），（0，1，0，1）。

我们将上述 10 个可取状态向量组成的集合记为 S，称 S 为允许状态集合。

2.状态转移方程

对于一次过河，可以看成一次状态转移，我们用向量来表示决策，例（1，0，0，1）表示狗，鸡过河。令 D 为允许决策集合，

$$D=\{(1, x, y, z): x+y+z=0 \text{ 或 } 1\}$$

另外，我们注意到过河有两种，奇数次的为从南岸到北岸，而偶数次的为从北岸回到南岸，因此得到下述转移方程，

$$S_{k+1} = S_k + (-1)^k d_k \tag{2.1}$$

$S_k = (w_k, x_k, y_k, z_k)$ 表示第 k 次状态，$d_k \in D$ 为决策向量。

人、狗、鸡、米过河问题，即要找到 $d_1, d_2, \cdots, d_{m-1} \in D$，$S_0, S_1, \cdots, S_m \in S$，$S_0 = (0,0,0,0)$，$S_m = (1,1,1,1)$ 且满足（2.1）式。

下面用状态转移图求解：

将 10 个允许状态用 10 个点表示，并且仅当某个允许状态经过一个允许决策仍为允许状态时，这两个允许状态间存在连线，而构成一个图， 如图 2-4。

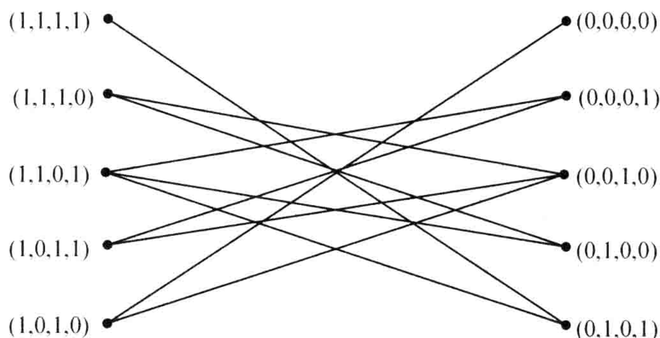


图 2-4

在其中寻找一条从（1，1，1，1）到（0，0，0，0）的路径，这样的路径就是一个解， 可得下述路径图：

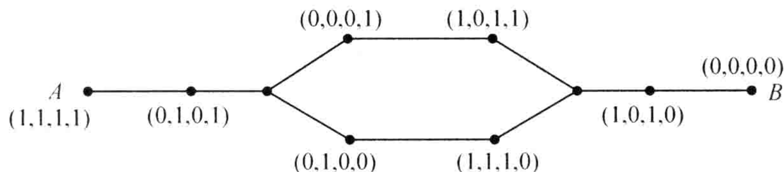


图 2-5

由图 2-5，有两个解都是经过 7 次运算完成，均为最优解。

2.2.2　商人过河问题

三名商人各带一个随从乘船渡河，现有一只小船，能容纳两个人，由他们自己划行，若在河的任一岸的随从人数多于商人，他们就可能抢劫财物。但如何乘船渡河由商人决定，试给出一个商人安全渡河的方案。

首先介绍图论中的一个定理：

G 是一个图，V（G）为 G 的顶点集，E（G）为 G 的边集。设 G 中有 n 个顶点 $v_1, v_2, \cdots, v_n$；$A=(a_{ij})_{n\times n}$ 为 G 的邻接距阵，其中

$$a_{ij}=\begin{cases}1, & v_iv_j\in E(G)\\ 0, & v_iv_j\notin E(G)\end{cases},\quad i,j=1,2,\cdots,n$$

定理 2.1 设$A(G)$为图G的邻接距阵，则G中从顶点v_i到顶点v_j，长度为k的道路的条数为A^k中的i行j列元素。

证明 对k用数学归纳法：

$k=1$时，显然结论成立；

假设k时定理成立，考虑 $k+1$的情形。

记A^l的i行j列元素为 $a_{ij}^{(l)}$，$l\geq 2$，因为$A^l\cdot A=A^{l+1}$，所以

$$a_{ij}^{(l+1)}=a_{i1}^{l}a_{1j}+a_{i2}^{l}a_{2j}+\cdots+a_{in}^{l}a_{nj} \tag{2.2}$$

而从 v_i到 v_j长$k+1$的道路无非是从v_i经k步到某顶点v_l ,$1\leq l\leq n$，再从v_l走一步到v_j；由归纳假设，从v_i到v_l长为k的道路共计a_{il}^{k}条，而从v_l到v_j长为 1 的道路为a_{lj}条，所以长为$k+1$的从v_i经k步到v_l再经一步到v_j的道路共有$a_{il}^{(k)}a_{lj}$条，故从v_i经$k+1$步到v_j的路径共有$a_{ij}^{(k+1)}=\sum_{l=1}^{n}a_{il}^{(k)}a_{lj}$条。

下面分析及求解：

假设渡河是从南岸到北岸，（m，n）表示南岸有 m 个商人，n 个随从，全部的允许状态共有 10 个：

$$v_1=(3,3),\ v_2=(3,2),\ v_3=(3,1),\ v_4=(3,0),\ v_5=(2,2)$$
$$v_6=(1,1),\ v_7=(0,3),\ v_8=(0,2),\ v_9=(0,1),\ v_{10}=(0,0)$$

以$V=\{v_1,v_2,\cdots,v_{10}\}$为顶点集，考虑到奇数次渡河及偶数次渡河的不同，我们建立两个邻接距阵

$$A=\begin{bmatrix}A_1 & A_2\\ 0 & A_3\end{bmatrix},\quad B=A^{\mathrm{T}}$$

其中

$$A_1=\begin{bmatrix}0&1&1&0&1\\0&0&1&1&1\\0&0&0&1&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix},\quad A_2=\begin{bmatrix}0&0&0&0&0\\0&0&0&0&0\\1&0&0&0&0\\0&0&0&0&0\\1&0&1&0&0\end{bmatrix},\quad A_3=\begin{bmatrix}0&0&0&1&1\\0&0&1&1&0\\0&0&0&1&1\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix}$$，其中A表示从南岸到北岸渡河的图的邻接距阵，$B=A^{\mathrm{T}}$表示从北岸到南岸渡河的图的邻接距阵。

由定理 2.1，我们应考虑最小的k，$(AB)^kA$ 中 1 行 10 列的元素不为 0，此时$2k+1$

即为最少的渡河次数，而矩阵 $(AB)^k A$ 中 1 行 10 列的元素为最佳的路径数目。

经过计算 k=5 时，$(AB)^5 A$ 的第 1 行 10 列元素为 2，所以需 11 次渡河，有两条最佳路径。

最后，我们用图解法求解（图 2-6）：

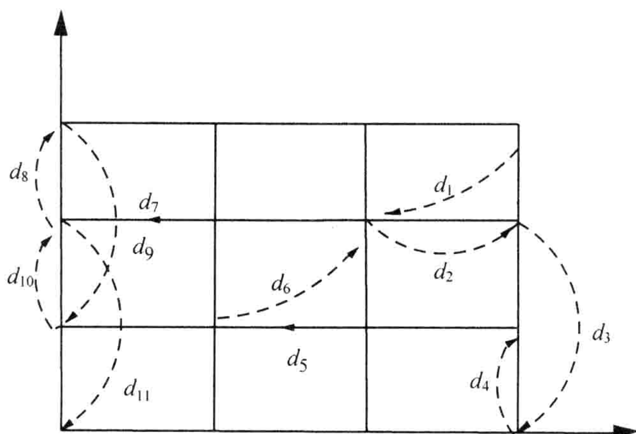


图 2-6

前面我们已求出问题的 10 种允许状态，允许决策向量集合 $D=\{(u, v): u+v=1, 2\}$，

状态转移方程为 $S_{k+1}=S_k+(-1)^k d_k$，如图 2-6，标出 10 种允许状态，找出从 d_1 经由允许状态到原点的路径，该路径还要满足奇数次向左、向下；偶数次向右、向上。

由图 2-6 可得这样的过河策略，共分 11 次决策：

$$(3,3)\xrightarrow{\text{去一商一随}}(2,2)\xrightarrow{\text{回一商}}(3,2)\xrightarrow{\text{去二随}}(3,0)\xrightarrow{\text{回一随}}(3,1)\xrightarrow{\text{去二商}}(1,1)$$

$$\xrightarrow{\text{回一商一随}}(2,2)\xrightarrow{\text{去二商}}(0,2)\xrightarrow{\text{回一随}}(0,3)\xrightarrow{\text{去二随}}(0,1)\xrightarrow{\text{回一随}}(0,2)$$

$$\xrightarrow{\text{去二随}}(0,0)$$

2.3 量纲分析法

量纲分析是 20 世纪初提出的，在物理领域中建立数学模型的一种方法，它是在经验和实验的基础上，利用物理定律的量纲齐次原则，确定各物理量之间的关系。

2.3.1 量纲齐次原则与 Pi 定理

许多物理量是有量纲的，有些物理量的量纲是基本的，另一些物理量的量纲则可以由基本量纲根据其定义或某些物理定律推导出来。例如在动力学中，把长度 l、质量 m 和时间 t 的量纲作为基本量纲，记为

$$[l]=L,\ [m]=M,\ [t]=T$$

而速度 v、力 f 的量纲可表示为

$$[v]=LT^{-1},\quad [f]=MLT^{-2}$$

在国际单位制中，有 7 个基本量：长度、质量、时间、电流、温度、光强度和物质的量，它们的量纲分别为 L、M、T、I、Θ、J 和 N，称为基本量纲。任一个物理量 q 的量纲都可以表示成基本量纲的幂次之积，

$$[q]=L^{\alpha}M^{\beta}T^{\gamma}I^{\delta}\Theta^{\varepsilon}N^{\xi}J^{\eta}$$

量纲齐次性原则 用数学公式表示一个物理定律时，等式两端必须保持量纲一致。

量纲分析就是在保证量纲一致的原则下，分析和探求物理量之间的关系。下面先看一个具体的例子，再给出量纲分析的一般方法。

例 2-1 单摆运动中，质量为 m 的小球系在长度为 l 的线的一端，线的另一端固定，小球偏离平衡位置后，在重力 mg 作用下做往复摆动，忽略阻力，求摆动周期 t 的表达式。

解 在这个问题中，有关的物理量有 t,m,l,g，设它们之间有关系式

$$t=\lambda m^{\alpha_1}l^{\alpha_2}g^{\alpha_3} \tag{2.3}$$

其中 α_1，α_2，α_3 为待定常数，λ 为无量纲的比例系数，则（2.3）式的量纲表达式为

$[t]=[m]^{\alpha_1}[l]^{\alpha_2}[g]^{\alpha_3}$，整理得

$$T=M^{\alpha_1}L^{\alpha_2+\alpha_3}T^{-2\alpha_3} \tag{2.4}$$

由量纲齐次原则应有

$$\begin{cases}\alpha_1=0\\ \alpha_2+\alpha_3=0\\ -2\alpha_3=1\end{cases} \tag{2.5}$$

解得：$\alpha_1=0$，$\alpha_2=\dfrac{1}{2}$，$\alpha_3=-\dfrac{1}{2}$，代入（2.3）得

$$t=\lambda\sqrt{\frac{l}{g}} \tag{2.6}$$

（2.6）式与单摆的周期公式是一致的。

下面我们给出用于量纲分析建模的 Buckingham Pi 定理。

定理 2 设 n 个物理量 $x_1,x_2,\cdots,x_n$ 之间存在一个函数关系

$$f(x_1,x_2,\cdots,x_n)=0 \tag{2.7}$$

$[x_1],\cdots,[x_m]$ 为基本量纲，$m\le n$。x_i 的量纲可表示为

$$[x_i]=\prod_{j=1}^{m}[x_j]^{\alpha_{ij}}\ ,\ i=1,2,\cdots,n$$

矩阵 $A=(\alpha_{ij})_{n\times m}$ 称为量纲距阵，若 $rankA=r$, 则（2.7）式与下式等价，

$$F(\pi_1,\pi_2,\cdots,\pi_{n-r})=0$$

其中 F 为一个未定的函数关系，π_s 为无量纲量，且 π_s 可表示为

$$\pi_s=\prod_{i=1}^{n}x_i^{\beta_i^{(s)}} \tag{2.8}$$

而 $\beta^{(s)}=(\beta_1^{(s)},\beta_2^{(s)},\cdots,\beta_n^{(s)})$ 为线性齐次方程组 $A^T\beta=0$ 的基本解向量。

利用 Pi 定理建模，关键是确定与该问题相关的几个基本量纲的无量纲量 $\pi_1,\pi_2,\cdots,\pi_{n-r}$。

作为量纲分析法的应用，下面我们介绍航船阻力的建模。

2.3.2 航船的阻力

长 l 、吃水深度 h 的船以速度 v 航行，若不考虑风的影响，航船受到的阻力 f 除依赖于船的诸变量 l、h、v 以外，还与水的参数——密度 ρ，黏性系数 μ 以及重力加速度 g 有关。

我们利用 Pi 定理分析 f 和上述物理量之间的关系。

1.航船问题中涉及的物理量及其量纲为

$$\begin{cases}[f]=LMT^{-2}\\ [l]=L\\ [h]=L\\ [v]=LT^{-1}\\ [\rho]=L^{-3}M\\ [\mu]=L^{-1}MT^{-1}\\ [g]=LT^{-2}\end{cases}$$

我们要寻求的关系式为

$$\phi(f,l,h,v,\rho,\mu,g)=0 \tag{2.9}$$

这些物理量中涉及的基本量纲为 L、M、T。

2.写出量纲矩阵

$$A^{\mathrm{T}}=\begin{bmatrix}1 & 1 & 1 & 1 & -3 & -1 & 1\\ 1 & 0 & 0 & 0 & 1 & 1 & 0\\ -2 & 0 & 0 & -1 & 0 & -1 & -2\end{bmatrix}\begin{matrix}L\\ M\\ T\end{matrix}\ ,$$

$$rank\ A=3\ 。$$

3.解齐次线性方程组 $A^{\mathrm{T}}\beta=0$ ，可得 $n-rank\ A=4$ 个基本解向量

$$\begin{cases} \beta^{(1)} = (0,1,-1,0,0,0,0)^T \\ \beta^{(2)} = (0,1,0,-2,0,0,1)^T \\ \beta^{(3)} = (0,1,0,1,1,-1,0)^T \\ \beta^{(4)} = (1,-2,0,-2,-1,0,0)^T \end{cases}$$

由（2.8）式，可给出 4 个无量纲量

$$\begin{cases} \pi_1 = lh^{-1} \\ \pi_2 = lv^{-2}g \\ \pi_3 = lv\rho\mu^{-1} \\ \pi_4 = fl^{-2}v^{-2}\rho^{-1} \end{cases} \quad (2.10)$$

由 Pi 定理，（2.9）等价于下列方程

$$\Phi(\pi_1,\pi_2,\pi_3,\pi_4) = 0 \quad (2..1)$$

这里 Φ 是未定的函数。

由（2.10），（2.11）可得阻力 f 的显式表达式，

$$f = l^2v^2\rho\psi(\pi_1,\pi_2,\pi_3) \quad (2.12)$$

其中 ψ 是由（2.11）可得到的未知的函数关系，在力学上，$\frac{v}{\sqrt{l\mathrm{g}}}$ 称为 Froude 数，记为 Fr；$\frac{lv\rho}{\mu}$ 称为 Reynold 数，记为 Re，因此（2.12）又可写为

$$f = l^2v^2\rho\psi(l/h, Fr, Re) \quad (2.13)$$

4.下面我们利用物理模拟进一步确定航船在水中的阻力

设 f、l、h、v、ρ、μ、g 和 f'、l'、h'、v'、ρ'、μ'、g' 分别表示模型和原型中的各物理量，由（2.13）有

$$f = l^2v^2\rho\psi\left(\frac{l}{h},\frac{v}{\sqrt{l\mathrm{g}}},\frac{lv\rho}{\mu}\right),\quad f' = l'^2v'^2\rho'\psi\left(\frac{l'}{h'},\frac{v'}{\sqrt{l'\mathrm{g}'}},\frac{l'v'\rho'}{\mu'}\right)$$

当无量纲量

$$\frac{l}{h} = \frac{l'}{h'},\quad \frac{v}{\sqrt{l\mathrm{g}}} = \frac{v'}{\sqrt{l'\mathrm{g}'}},\quad \frac{lv\rho}{\mu} = \frac{l'v'\rho'}{\mu'} \quad (2.14)$$

成立时，可得 $\frac{f'}{f} = \left(\frac{l'v'}{lv}\right)^2\frac{\rho'}{\rho}$。 (2.15)

则此时由模型船的阻力 f 及其他的 l、v、ρ、l'、v'、ρ' 可确定原型船的阻力 f'。

下面，我们讨论一下（2.14）成立的条件，如果在实验中采用跟实际同样的水质，则 $\rho' = \rho, \mu' = \mu$；又 $g = g'$，故可得

$$\frac{l}{h}=\frac{l'}{h'}, \frac{v}{v'}=\sqrt{\frac{l}{l'}}, \frac{v}{v'}=\frac{l'}{l} \tag{2.16}$$

要使得（2.16）成立，必有 $l=l', h=h'$；也即模型船与原型船一样大，这显然排除了物理模拟的可行性。若考虑选用不同的水质，$\mu'\neq\mu$，仍设 $\rho'=\rho$，则（2.14）式化为

$$\frac{l}{h}=\frac{l'}{h'}, \frac{v}{v'}=\sqrt{\frac{l}{l'}}, \frac{v}{v'}=\frac{l'}{l}\cdot\frac{\mu}{\mu'} \tag{2.17}$$

由（2.17）可得 $\frac{\mu}{\mu'}=\left(\frac{l}{l'}\right)^{\frac{3}{2}}$，若按 1∶20 的比例，$\mu=0.011\mu'$，显然无法找到如此小的黏性系数的液体。实际上的一种近似处理方法是，在一定条件下 Re 数的影响很小，这样可近似得到，

$$f\approx l^2v^2\rho\psi(\frac{l}{h}, \frac{v}{\sqrt{lg}})$$

类似地分析，只要 $\frac{l}{h}=\frac{l'}{h'}$，$\frac{v}{v'}=\sqrt{\frac{l}{l'}}$ 即有

$$\frac{f'}{f}=\left(\frac{l'}{l}\right)^3 \tag{2.18}$$

由（2.18）式就容易确定原型船的阻力 f'。

2.3.3　无量纲化抛射问题

下面我们通过一个例子，介绍如何使用无量纲化方法简化模型。

抛射问题　在某星球表面以初速度 v 竖直向上发射火箭，记星球半径为 r，星球表面重力加速度为 g，忽略阻力，讨论发射高度 x 随时间 t 的变化规律。

模型建立　设 x 轴竖直向上，$t=0$ 时 $x=0$，火箭和星球质量分别记为 m_1 和 m_2，由牛顿第二定律和万有引力定律可得:

$$m_1\ddot{x}=-k\frac{m_1m_2}{(x+r)^2} \tag{2.19}$$

以 $x(0)=0, \ddot{x}(0)=-g$ 代入（2.19）可得 $km_2=gr^2$，所以，得如下初值问题

$$\begin{cases}\ddot{x}=-\dfrac{r^2g}{(x+y)^2}\\ x(0)=0\\ \dot{x}(0)=v\end{cases} \tag{2.20}$$

（2.20）式的解可以表示为 $x=(t, r, v, g)$，即发射高度是以 r、v、g 为参

数的t的函数，下面我们采用无量纲化方法化简方程（2.20）。

显然，抛射问题中的基本量纲为 L，T；而

$$[x]=L,[t]=T,[r]=L,[v]=LT^{-1},[g]=LT^{-2}\text{。}$$

所谓无量纲化是指，对（2.20）式中的x和t分别构成且有相同的参数组合x_c和t_c，使得新变量 $\bar{x}=\dfrac{x}{x_0}$，$\bar{t}=\dfrac{t}{t_0}$为无量纲量，其中x_c,t_c 称为特征尺度或参考尺度；把方程（2.20）化为$\bar{x}$对$\bar{t}$的微分方程，即可简化模型。

如何寻找特征尺度？这里我们以t_c为例，首先写出参数r、v、g的量纲距阵A，

$$A=\begin{bmatrix}1 & 1 & 1\\0 & -1 & -2\end{bmatrix}$$

t的量纲向量为 $(0,1)^{\mathrm{T}}$，记为β_0。

求解线性方程组 $A\beta=\beta_0$得通解

$$\beta=(1,-1,0)^{\mathrm{T}}+k(1,-2,1)^{\mathrm{T}}$$

任取k，即得到一种特征尺度，例如$k=0$得$t_c=rv^{-1}$；$k=-1$得$t_c=vg^{-1}$；$k=-\dfrac{1}{2}$得$t_c=\sqrt{rg^{-1}}$。

同理可得x的几种特征尺度r,v^2g^{-1}等。

以下我们利用不同的x_c和t_c化简（2.20）。

1.令$x_c=r,t_c=rv^{-1}$，则$\bar{x}=\dfrac{x}{r},\bar{t}=\dfrac{t}{rv^{-1}}$，将$\dot{x}=v\dfrac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{t}}$，$\ddot{x}=\dfrac{v^2}{r}\dfrac{\mathrm{d}^2\bar{x}}{\mathrm{d}\bar{t}^2}$

代入（2.20）可得:

$$\begin{cases}\varepsilon\ddot{\bar{x}}=-\dfrac{1}{(\bar{x}+1)^2}\text{，}\varepsilon=\dfrac{v^2}{rg}\\ \bar{x}(0)=0\\ \dot{\bar{x}}(0)=1\end{cases}\tag{2.21}$$

（2.21）式的解可表为 $\bar{x}=\bar{x}(\bar{t},\varepsilon)$，含一个独立参数且$\varepsilon$为无量纲量。

2. 令$x_c=r$，$t_c=\sqrt{rg^{-1}}$， 类似地可将（2.20）化为

$$\begin{cases}\ddot{\bar{x}}=-\dfrac{1}{(\bar{x}+1)^2}\\ \bar{x}(0)=0\\ \dot{\bar{x}}(0)=\sqrt{\varepsilon}\text{，}\varepsilon=\dfrac{v^2}{rg}\end{cases}\tag{2.22}$$

3. 令 $x_0 = v^2 g^{-1}$，$t_0 = vg^{-1}$，可将（2.21）化为

$$\begin{cases} \varepsilon\ddot{\bar{x}} = -\dfrac{1}{(\varepsilon\bar{x}+1)^2}，\ \varepsilon = \dfrac{v^2}{rg} \\ \bar{x}(0) = 0 \\ \dot{\bar{x}}(0) = 1 \end{cases} \tag{2.23}$$

按照现有科技能力，$v << \sqrt{rg} \approx 8000$米/秒，$\varepsilon << 1$。在（2.23）中令 $\varepsilon = 0$，则有

$$\begin{cases} \ddot{\bar{x}} = -1 \\ \bar{x}(0) = 0 \\ \dot{\bar{x}}(0) = 1 \end{cases} \tag{2.24}$$

（2.24）的解为：$\bar{x}(\bar{t}) = -\dfrac{\bar{t}^2}{2} + \bar{t}$，代回原变量得

$$x(t) = -\frac{1}{2}gt^2 + vt \tag{2.25}$$

（2.25）式恰为假定火箭运动过程中所受星球引力不变的运动方程。

小结　无量纲化是用数学工具研究物理问题时常用的方法，恰当地选择特征尺度不仅可以减少参数的个数，而且可以帮助人们决定舍弃哪些次要因素。

2.4 比例与函数建模

本节介绍的几个模型，都是利用基本的比例关系与函数建立起的数学模型。

2.4.1 动物体型问题

问题　某生猪收购站，需要研究如何根据生猪的体长（不包括头尾）估计其体重。

模型假设

1.将四足动物的躯干（不含头尾）视为质量为 m 的圆柱体，长度为 l，截面面积 S，直径为 d，如图 2-7。

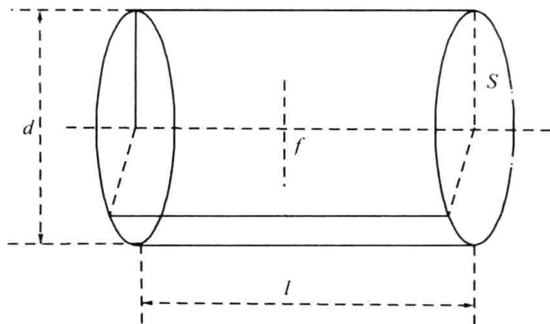


图 2-7

2.把圆柱体的躯干看作一根支撑在四肢上的弹性梁，动物在体重 f 作用下的最大下垂为 δ，即梁的最大弯曲，根据弹性力学弯曲度理论，有

$$\delta \propto \frac{fl^3}{sd^2} \tag{2.26}$$

3.以生物进化学的角度，可认为动物的相对下垂度 $\frac{\delta}{l}$ 已达到一个最合适的数值，也即 $\frac{\delta}{l}$ 为常数。

模型建立

因为 $S=\frac{\pi d^2}{4}$，所以

$$f=\frac{\pi d^2}{4}l \tag{2.27}$$

由（2.26）式可令 $\delta = k_1\frac{fl^3}{Sd^2}$，$k_1$ 为比例常数，由（2.27）式得

$$Sd^2=\frac{4}{\pi}(\frac{\pi}{4}d^2l)^2\cdot\frac{1}{l^2}=\frac{4}{\pi}f^2\cdot\frac{1}{l^2} \tag{2.28}$$

所以 $\delta = k_1\frac{\pi}{4}\frac{l^5}{f}$，于是 $f=\frac{\pi}{4}k_1\frac{l^5}{\delta}=\frac{\pi}{4}k_1\frac{l}{\delta}\cdot l^4$。

令 $k=\frac{\pi}{4}k_1\frac{l}{\delta}$，由假设 3，$k$ 为常数，所以

$$f=kl^4 \tag{2.29}$$

因此，生猪的体重与体长的四次方成正比，在实际工作中，工作人员可由实际经验及统计数据找出常数 k，则可近似地由生猪的体长估计它的体重。

2.4.2 双重玻璃的功效

问题 房间的窗户有的是双层的，即在窗户上装两层玻璃，且中间留有一定的空隙，试比较双层玻璃窗与单层玻璃窗的热量流失。

模型假设

1.设双层玻璃窗的两玻璃的厚度都为 d，两玻璃的间距为 L；单层玻璃窗的玻璃厚度为 $2d$，所用玻璃材料相同，如图 2-8。

2.假设窗户的封闭性能很好，两层玻璃之间的空气不流动，即忽略热量的对流，只考虑热量的传导。

3.室内温度 T_1 和室外温度 T_2 保持不变，热传导过程处于稳定状态，即单位时间通过单位面积的热量为常数。

4.玻璃材料均匀，热传导系数为常数。

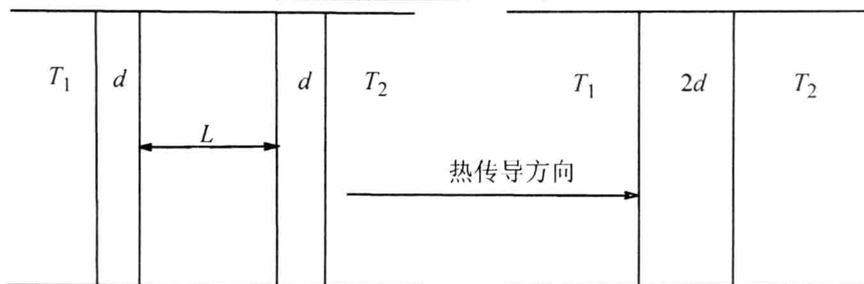


图 2-8

模型建立

对于厚度为 d 的均匀介质，两侧温度差为 ΔT，则单位时间内由温度高的一侧向温度低的一侧通过单位面积的热量 Q 满足 $Q = k \cdot \dfrac{\Delta T}{d}$，$k$ 为热传导系数。

设玻璃的热传导系数为 k_1，空气的热传导系数为 k_2。

（1）先考虑单层玻璃的单位时间、单位面积的热量传导。

$$Q_1 = k_1 \cdot \frac{T_1 - T_2}{2d} \tag{2.30}$$

（2）考虑双层玻璃情形

此时热量先通过厚度为 d 的玻璃传导到两层玻璃的夹层空气中，再通过空气传导，再通过厚度为 d 的玻璃传导；设内层玻璃的外侧温度为 T_a，外层玻璃的内侧温度为 T_b，则有:

$$Q_2 = k_1 \frac{T_1 - T_a}{d} = k_2 \frac{T_a - T_b}{l} = k_1 \frac{T_b - T_2}{d} \tag{2.31}$$

由（2.31）式可得

$$\begin{cases} T_a + T_b = T_1 + T_2 \\ T_a - T_b = \dfrac{k_1}{k_2} \dfrac{l}{d} (T_b - T_2) \end{cases} \tag{2.32}$$

记 $s = \dfrac{k_1 l}{k_1 d}$，则有

$$2T_b = T_1 + T_2 - s(T_b - T_2)\text{；} \tag{2.32}$$

$$2(T_b - T_2) = T_1 - T_2 - s(T_b - T_2)\text{；} \tag{2.33}$$

$$(T_b - T_2) = \frac{1}{2+s}(T_1 - T_2)\text{；} \tag{2.34}$$

$$Q_2 = \frac{1}{(2+s)} \frac{k_1}{d} (T_1 - T_2)\text{；} \tag{2.35}$$

考虑两者之比

$$\frac{Q_2}{Q_1}=\frac{2}{2+s} \qquad (2.36)$$

显然 $Q_2<Q_1$， 也即双层玻璃的热量损失较小。

模型分析与应用

常用玻璃的热传导系数 $k_1=4\times10^{-3}\sim8\times10^{-3}$J/（cm·s·℃），而不流通、干燥空气的热传导系数 $k_2=2.5\times10^{-4}$J/（cm·s·℃）。

若取 $\frac{l}{d}=h$，则 $16h\le s\le 32h$ ，故 $\frac{Q_2}{Q_1}\le\frac{1}{1+8h}$ （2.37）

若取 $h=4$， 则 $\frac{Q_2}{Q_1}\le\frac{1}{33}$，由此可见双层玻璃的保暖效果是相当可观的。

我国北方寒冷地区的建筑物，通常采用双层玻璃；由（2.37）式，h=4 时，$Q_2\approx\frac{1}{33}Q_1$ ， h 再大，热量传递的减少就不明显了，再考虑到墙体的厚度；所以建筑规范通常要求 $h\approx4$。

2.4.3 席位分配模型

问题 某校有 200 名学生，甲系 100 名，乙系 60 名，丙系 40 名；若学生会中学生代表有 20 个席位，则公平又简单的分法应各有 10、6、4 个席位。若丙系有 6 名学生分别转入甲、乙两系各 3 人，此时各系的人数为 103、63、34；按比例席位分配应为 10.3，6.3 和 3.4，出现了小数，19 个整数席位分配完后，最后一席留给小数部分最大的丙系，分别为 10，6，4。为方便提案表决，现增加 1 席共 21 席，按比例计算甲、乙、丙三系分别占有 10.815，6.615，3.570；按上面的分法应分别为 11，7，3；这样虽然增加了一个席位，但丙系的席位反而减少一席，因此这种分法显然是不合理的，请给出一个比较公平的席位分配方案。

问题分析

席位分配问题，当出现小数时，无论如何分配都是不完全公平的。那么一个比较公平的分法是应该找到一个不公平程度最低的方法，因此首先要给出不公平程度的数量化，然后考虑使之最小的分配方案。

模型建立

1.讨论不公平程度的数量化。设 A、B 两方人数分别为 p_1，p_2；分别占有 n_1 和 n_2 个席位，则两方每个席位所代表的人数分别为 $\frac{p_1}{n_1}$ 和 $\frac{p_2}{n_2}$。我们称 $\left|\frac{p_1}{n_1}-\frac{p_2}{n_2}\right|$ 为绝对不公平值。

例 $p_1=120$，$p_2=100$，$n_1=n_2=10$，则$\left|\frac{p_1}{n_1}-\frac{p_2}{n_2}\right|=2$，

又 $p_1=1020$，$p_2=1000$，$n_1=n_2=10$，则 $\left|\frac{p_1}{n_1}-\frac{p_2}{n_2}\right|=2$。

由上例可知，用绝对不公平程度作为衡量不公平的标准，并不合理，下面我们给出相对不公平值。

若$\frac{p_1}{n_1}>\frac{p_2}{n_2}$，则称$\dfrac{\frac{p_1}{n_1}-\frac{p_2}{n_2}}{\frac{p_2}{n_2}}=\frac{p_1 n_2}{p_2 n_1}-1$为对 A 的相对不公平值，记为 $r_A(n_1,\ n_2)$。

若$\frac{p_1}{n_1}<\frac{p_2}{n_2}$，则称$\dfrac{\frac{p_2}{n_2}-\frac{p_1}{n_1}}{\frac{p_1}{n_1}}=\frac{p_2 n_1}{p_1 n_2}-1$为对 B 的相对不公平值 ，记为 $r_B(n_1,\ n_2)$。

上例中，相对不公平值分别为 0.2 和 0.02，可见相对不公平值较合理。

2.用相对不公平值建立模型

设 A、B 两方人数分别为 p_1，p_2，分别占有 n_1 和 n_2 个席位。现在增加一个席位，应该给 A 还是 B ？不妨设$\frac{p_1}{n_1}>\frac{p_2}{n_2}$ ，此时对 A 不公平，下面分两种情形。

（1）$\frac{p_1}{n_1+1}\geq\frac{p_2}{n_2}$，这说明即使 A 增加 1 席，仍对 A 不公平， 故这一席应给 A。

（2）$\frac{p_1}{n_1+1}<\frac{p_2}{n_2}$，说明 A 方增加 1 席时，将对 B 不公平，此时计算对 B 的相对不公平值

$$r_B(n_1+1,\ n_2)=\frac{p_2(n_1+1)}{p_1 n_2}-1$$

若这一席给 B，则对 A 的相对不公平值为

$$r_A(n_1,\ n_2+1)=\frac{p_1(n_2+1)}{p_2 n_1}-1$$

本着使得相对不公平值尽量小的原则，若

$$r_B(n_1+1,\ n_2)<r_A(n_1,\ n_2+1) \tag{2.38}$$

则增加的 1 席给 A 方，若

$$r_A(n_1,\ n_2+1)<r_B(n_1+1,\ n_2) \tag{2.39}$$

则增加的 1 席给 B 方。由（2.38）式可得：$\frac{p_2^2}{n_2(n_2+1)}<\frac{p_1^2}{n_1(n_1+1)}$ ；由（2.39）式

可得：$\dfrac{p_2^2}{n_2(n_2+1)} > \dfrac{p_1^2}{n_1(n_1+1)}$。

记：$Q_i = \dfrac{p_i}{n_i(n_i+1)}$，则增加的1席，应给$Q$值大的一方。

第一种情形，显然也符合该原则。

现在将上述方法推广到m方分配席位的情况，A_i方人数为p_i，已占有n_i席，$i=1,2,\cdots,m$。

计算$Q_i = \dfrac{p_i^2}{n_i(n_i+1)}$，则应将增加的1席分配给$Q$值最大的一方。

下面考虑原问题：

前19席的分配没有争议，甲系得10席，乙系得6席，丙系得3席。

第20席的分配：

$$Q_1 = \frac{103^2}{10(10+1)} = 96.4$$

$$Q_2 = \frac{63^2}{6(6+1)} = 94.5$$

$$Q_3 = \frac{34^2}{3(3+1)} = 96.3$$

故第20席分配给甲系。

第21席的分配：

$$Q_1 = \frac{103^2}{11(11+1)} = 80.4$$

$$Q_2 = 94.5\ ,\ Q_3 = 96.3$$

故第21席分配给乙系。

甲、乙、丙三系各分得11，6，4席，这样丙系保住它险些丧失的1席。

习题二

1.设一所监狱有64间囚室，其排列类似8×8棋盘，看守长告诉关押在一个角落里的囚犯，只要他能够不重复地通过每间囚室到达对角的囚室（所有相邻囚室间都有门相通），他将获释，问囚犯能否获得自由?

2.某班有49个学生，坐成7行7列，每个坐位的前后左右的坐位叫做它的邻座，要让49个学生都换到他的邻座上去，问是否有这种调换位置的方案?

3.9个人的集会中一定有3个人互相认识或4个人互相不认识，对吗?

4.14 个人的集会中一定有 3 个人互相认识或者 5 个人互相不认识，对吗?

5.17 个科学家中，每个科学家都和其他科学家通信，他们之间讨论 3 个题目，且任意两个科学家之间只讨论 1 个题目，求证：其中至少有 3 名科学家，他们互相通信中讨论的是同一题目。

6.夫妻过河问题：有 3 对夫妻过河，船最多能载 2 人，条件是任一女子不能在其丈夫不在的情况下与其他男子在一起，如何安排三对夫妻过河? 若船最多能载 3 人，5 对夫妻能否过河?

7.怎样讨论三层玻璃的隔热功效?

8.怎样讨论双层玻璃的隔音效果?

第3章　线性代数建模

向量、向量空间、矩阵等都是线性代数中的重要概念，本章将通过一些简单的实例来说明它们在实际中的应用。

3.1　几个数学游戏

3.1.1　（人、狗、鸡、米过河问题）

这是一个人所共知而又十分简单的智力游戏。某人要带狗、鸡、米过河，但小船除了需要有人去划以外，最多只能载一物过河，而当人不在场时，狗要咬鸡、鸡要吃米，问此人应如何过河。

要知道例 3.1 的答案并不困难。第一次，人只能带鸡过河。到了对岸，人只有自己回来，将鸡留在对岸，否则，又返回了初始状态。接下来，人可以带狗过河，也可以带米过河，但回来时一定要将鸡带回……按此推导下去，读者不难找到过河的方法。

我们研究本例的目的不在于找出答案，而是想设计出一种让计算机自行搜索寻找答案的方法。为达到此目的，我们先把比例转化为状态转移问题。首先，应当如何表达状态呢？不同的情况应采取不同的方法，在本例中，人、鸡、狗、米都只有两种可能状态，即在此岸或在彼岸（不在此岸）。我们用向量表示状态，可采取如下方法：一物在此岸时相应分量取 1，而在彼岸时则相应分量取 0，例如（1,0,1,0）表示人和鸡在此岸，而狗和米则在彼岸。

（1）可取状态：根据题意，并非所有状态都是允许的，例如（0,1,1,0）就是一个不可取的状态，因为狗会咬鸡。本题中可取状态（即系统允许的状态）可以用穷举法列出来，它们是：

人在此岸	人在对岸
(1，1，1，1)	(0，0，0，0)
(1，1，1，0)	(0，0，0，1)
(1，1，0，1)	(0，0，1，0)
(1，0，1，1)	(0，1，0，0)
(1，0，1，0)	(0，1，0，1)

总共有十个可取状态。对于一般情况，也可找出状态为可取的充要条件，让计算机根据充要条件来检查得到的状态是否为可取状态。

（2）可取运算：状态转移需要经过状态运算来实现。在实际问题中，摆一次渡即可改变现有状态。为此再引入一个四维向量（转移向量），用它来反映摆渡情况。例如（1，1，0，0）表示人带狗摆渡过河。根据题意，允许使用的转移向量只能有（1，0，0，0，）、（1，1，0，0）、（1，0，1，0）、（1，0，0，1）四个。

为实现本题中的状态转移，规定一个状态向量与转移向量之间的运算。规定状态向量与转移向量之和为一新的状态向量，其运算为对应分量相加，且规定 0+0=0，1+0=0+1=1，1+1=0。例如（1，1，1，1）+（1，0，1，0）=（0，1，0，1），其实际意义为人、狗、鸡、米原来均在此岸，人带鸡过河，转变为新的状态，此岸仅剩狗和米（注：此处作这样的运算规定完全是为了与实际情况相符）。

在具体转移时，只考虑由可取状态到可取状态的转移。问题化为：由初始状态（1，1，1，1）出发，经过奇数次的上述运算能否转化为（0，0，0，0）的转化问题，进而，如果能，我们还想知道具体应当如何转化。

由于规定的运算十分容易在计算机上实现，这样一来就把一个数学游戏转化为了一个可以在计算机上计算的数学问题（即建模）。当然，像本题这样简单的问题，也可通过笔算方法求解，具体可如下进行分析。

（第一次渡河）

$$(1,1,1,1)+\begin{cases}(1,1,0,0)\\(1,0,1,0)\\(1,0,0,1)\\(1,0,0,0)\end{cases}=\begin{cases}(0,0,1,1) & \times & (\text{不可取})\\(0,1,0,1) & & (\text{可取})\\(0,1,1,0) & \times & (\text{不可取})\\(0,1,1,1) & \times & (\text{不可取})\end{cases}$$

（第二次渡河）

$$(0,1,0,1)+\begin{cases}(1,1,0,0)\\(1,0,1,0)\\(1,0,0,1)\\(1,0,0,0)\end{cases}=\begin{cases}(1,0,0,1) & \times & (\text{不可取})\\(1,1,1,1) & \times & (\text{循环,回到原先出现过的状态})\\(1,1,0,0) & \times & (\text{不可取})\\(1,1,0,1) & & (\text{可取})\end{cases}$$

以下可继续进行下去，直至转移目的实现。

上述分析实际上采用的是穷举法，对于规模较大的问题是不宜采用的。

3.1.2 夫妻过河问题

这是一个古老的阿拉伯数学问题。有三对夫妻要过河，船最多可载两人，约束条件是根据阿拉伯法律，任一女子不得在其丈夫不在场的情况下与其他男子在一起，问此时这三对夫妻能否过河？

这一问题的状态和运算与前一问题有所不同，根据题意，状态应能反映出两岸的男女人数，过河也同样要反映出性别，故可如下定义：

（1）可取状态：用 H 和 W 分别表示此岸的男子数和女子数，状态可用矢量（H, W）表示，其中 $0 \leqslant H$、$W \leqslant 3$。可取状态为 $(0,i),(i,i),(3,i)$，$0 \leq i \leq 3$。(i,i) 为可取状态，这是因为总可以适当安排而使他们是 i 对夫妻。

（2）可取运算：过河方式可以是一对夫妻、两个男人或两个女人，当然也可以是一人过河。转移向量可取成 $\left((-1)^i m,\ (-1)^i n\right)$，其中 m、n 可取 0、1、2，但必须满足 $1 \leqslant m+n \leqslant 2$。当 i 为奇数时表示过河；当 i 为偶数时表示由对岸回来，运算规则同普通向量的加法。问题归结为由状态（3, 3）经奇数次可取运算，即由可取状态到可取状态的转移，转化为（0, 0）的转移问题。和上题一样，我们既可以用计算机求解，也可以分析求解，此外，本题还可用作图方法来求解。

在 H—W 平面坐标中，以“·”表示可取状态，从 $A(3, 3)$ 经奇数次转移到达 $O(0, 0)$。奇数次转移时向左或下移动 1~2 格而落在一个可取状态上，偶数次转移时向右或上移动 1~2 格而落在一个可取状态上。另外，由于奇数次与偶数次过河产生的效果是不同的，为了区分起见，用实箭线表示奇数次转移，用虚箭线表示偶数次转移，图 3-1 给出了一种可实现的方案，故这三对夫妻是可以过河的。假如按这样的方案过河，共需经过 11 次摆渡。

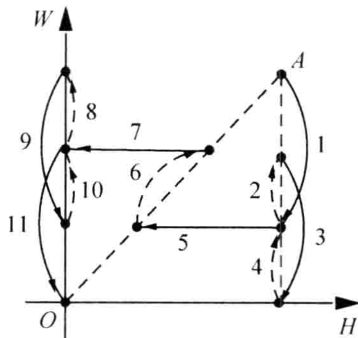


图 3-1

不难看出，在上述规则下，4 对夫妻就无法过河了，读者可以自行证明之。类似可以讨论船每次可载 3 人的情况，其结果是 5 对夫妻是可以过河的，而 6 对以上时就无法过河了。假如船每次可以载 4 人，则任意多对夫妻均可过河，最易看出的一个方案是让一对夫妻当船员工即可。

关于夫妻过河还可以编出许多其他形式的问题，下面我们来讨论一些同样有趣的问题。为了叙述简便，先约定一些符号，这些符号将被应用于以下的各个问题之中。记想过河的夫妻对数为 n，船可载的人数为 m，n 对夫妻过河所需的最少摆渡次数为 k。

问题 1　2 对夫妻要过河，船每次只能渡 2 人，应如何过河，最少摆渡几次？（即 $n=2$，$m=2$，求 k。）

本问题很容易解答，读者可自行完成（答案为 $k=5$）。

问题 2　n 对夫妻要过河，船每次可载 n-1 人，应如何过河，最少要摆几次渡？（$m=n-1$，求 k。）

答案如下：（1）$n=3$，$m=2$，$k=11$；

（2）$n=4$，$m=3$，$k=9$；

（3）$n\geq5$，$m=n-1$，$k=7$。

问题 3　1883 年，吕卡斯（Récréations）提出以下问题：n 对夫妻要过河，船至少应可载几人（$m\geq$?）他们才可能过河，最少摆渡次数为多少？

德兰努瓦（M. Delannoy）给出证明：

（1）$n=2$，$m=2$，$k=5$；

（2）$n=3$，$m=2$，$k=11$；

（3）$n=4$，$m=3$，$k=9$；

（4）$n=5$，$m=3$，$k=11$；

（5）$n\geq6$，$m=4$，　$k=2n-3$。

问题 4　德丰特内（M. de Fonteney）指出，如果河中有一个岛，那么，不管有多少对夫妻，只要有一只可载 2 人的船，他们均能过河（2 对、3 对时不需要岛），最少摆渡次数为 $8n-6$。

上面介绍的几个例子本身并无多大实际意义，但它们展示了如何将实际问题转化为状态转移问题的方法，这种方法是值得借鉴的。

3.2　Dürer 魔方（或幻方）问题

有些较为复杂的问题，开始时常常给人以一种变幻莫测的感觉。但经过细微的

分析研究，可以发现其中存在着某些内在的关系。在使用适当的数学工具后，这些内在关系就被一一揭露出来了。

德国著名的艺术家 Albrecht Dürer(1471-1521)于 1514 年曾铸造了一枚名为“Melencotia I”的铜币。令人奇怪的是在这枚铜币的画面上充满了数学符号、数字及几何图形。这里，我们仅研究铜币右上角的数字问题。

3.2.1 Dürer 魔方

这是一个由自然数组成的方块，称为 Dürer 魔方，其数字排列如下：

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}$$

什么是魔方？我们来下一个定义。所谓的魔方是指由 $1 \sim n^2$ 这 n^2 个正整数按一定规则排列成的一个 n 行 n 列的正方形。按不同的要求，它可以具有某些特定的性质，n 称为此魔方的阶。例如，上面给出的 Dürer 魔方是 4 阶的，它的每一行数字之和为 34，每一列数字之和为 34，把对角线（或反对角线）上的数字加起来是 34，每个小方块中的数字之和也是 34，若把四个角上的数字加起来还是 34，多么奇妙！最后一行中间两个数字恰好是铜币的铸造时间——1514 年。

构造魔方是一个古老的数学游戏，起初它还和神灵联系在一起，带有深厚的迷信色彩。传说 3200 多年前，因治水出名的皇帝大禹就构造了三阶魔方（被人们称“洛书”），

$$\begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} \quad \text{（被人称为洛书的 3 阶魔方）}$$

至今还有人把它当作符咒用于某些迷信活动。

大约在 15 世纪时，魔方传到了西方，著名的科尼利厄斯·阿格里帕（1486—1535）先后构造出了 3~9 阶的魔方。

如何构造出各种阶数的魔方呢？假如你知道方法，构造它其实并不困难。

在构造 n 阶魔方时，首先要看清 n 是奇数还是偶数，在构造时要巧妙地利用某种形式的对称性。

我们先来看 n 是奇数的情况，奇数阶魔方的构造方法如下：

首先，在第一行中间写 1；然后每次向右上方移一格，依次填由小到大排列的下一个数（注：向上移出界时填下一列最后一行的小方格；向右移出界时填第一列

上一行的小方格，就好像上下边是相连的、左右边也是相连的一样）。此外，当下面想填的格已填过数或已达到魔方的右上角时，改填刚才填的格子正下方的小方格，此后继续按原方法填，直至完全填完所有小方格。例如，按上述方法可构造出下面的 5 阶魔方：

$$\begin{bmatrix} 17 & 24 & 1 & 8 & 15 \\ 23 & 5 & 7 & 14 & 16 \\ 4 & 6 & 13 & 20 & 22 \\ 10 & 12 & 19 & 21 & 3 \\ 11 & 18 & 25 & 2 & 9 \end{bmatrix}$$

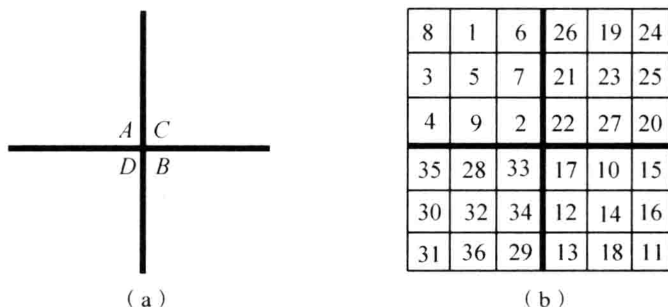
作为练习，请你给出一个 7 阶的魔方（见习题）。

偶数阶的魔方可以利用奇数阶魔方拼接而成，拉尔夫·斯特雷奇给出了一种拼接的方法。限于篇幅，我们不在此详细介绍了，作为一个例子，我们采用他的方法构造一个 6 阶的魔方。

第一步　利用 1~9，10~18，19~27 及 28~36 构造出 4 个 3 阶的魔方，它们分别是：

$$\begin{bmatrix} 8 & 1 & 5 \\ 3 & 4 & 7 \\ 6 & 9 & 2 \end{bmatrix} \quad \begin{bmatrix} 17 & 10 & 15 \\ 12 & 14 & 16 \\ 13 & 18 & 11 \end{bmatrix} \quad \begin{bmatrix} 26 & 19 & 24 \\ 21 & 23 & 25 \\ 22 & 27 & 20 \end{bmatrix} \quad \begin{bmatrix} 35 & 28 & 33 \\ 30 & 32 & 34 \\ 31 & 36 & 29 \end{bmatrix}$$

第二步　利用图 3-2(a)中的 A、B、C、D 容易拼出一个 6 阶的魔方。为了保证性质的成立，还需要作一些调整，如果你有兴趣，不妨找一下调整的方法，调整后得到的 6 阶魔方如图 3-2(b)所示。



8	1	6	26	19	24
3	5	7	21	23	25
4	9	2	22	27	20
35	28	33	17	10	15
30	32	34	12	14	16
31	36	29	13	18	11

（a）　　（b）

图 3-2

上述方法并非构造魔方的唯一方法，但不论采用什么方法来构造魔方，都应当尽可能利用某种形式的对称性。在魔方的构造中包含了许多有趣的数学问题，但由于很多人研究过这些问题，我们一般只能把它们当成一些练习题。

互不相同的同阶魔方究竟有多少个？人们知道，三阶魔方只有一个，当然，通

过镜面反射和绕中心旋转可以产生 8 种不同的表现形式。四阶魔方共有 880 个，而通过反射与旋转可有 7040 种不同的形式。没有人知道五阶或更高阶魔方的数量。例如，对五阶魔方，人们可用某种办法作出实质上不同的 57600 个（不含反射与旋转而得出的），如加上用其他方法构造的，已知的五阶魔方总数远远超过了 1300 万个，魔方数量随阶数 n 增长已达到了惊人的速度，令人目瞪口呆。

3.2.2 松弛问题的讨论

假如我们放松对构造魔方的数必须是 $1\sim n^2$ 的要求而允许它们取任意实数（就像将整数规划或 0-1 规划松弛成相应的线性规划那样），问题也许会简单得多。我们仍要求它们具有某些特定的性质，并不妨仍把它们称为魔方，当然，此时问题已发生了实质性的变化，不再是原先讨论的问题了。

为简单起见，我们用 n 阶方阵来记这样的魔方。易见，若 A 与 B 均为具有指定性质的魔方，则对任意的实数 α 和 β，$\alpha A+\beta B$ 也是。这一性质表明，具有指定性质的魔方全体构成一个线性空间。根据线性代数知识，要刻画一个线性空间只需指出它的维数并求出此线性空间的一组基底即可。

不妨仍以 4 阶方阵为例。首先，定义 0 方与 1 方如下：

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$R=C=D=S=0 \qquad R=C=D=S=4$$

其中 R 为行和，C 为列和，D 为对角线和，S 为小方块和。

现在，我们来研究具有性质 $R=C=D=S$ 的方阵构成的线性空间 X，类似于构造 n 维欧氏空间的标准基，我们利用 0 和 1 来构造一些 $R=C=D=S=1$ 的最简单的方阵，不难看出，1 在第一行中共有 4 种排法，为保持上述性质的成立，在第一行的 1 取定后，第二行中的 1 尚有两种取法。当第二行的 1 也取定后，第三行与第四行的 1 就完全定位了，故一共可作出 8 个不同的最简方阵，称之为基本魔方，并记之为 Q_1，…，Q_8。

$$Q_1=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad Q_2=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$Q_3=\begin{bmatrix}0&0&0&1\\1&0&0&0\\0&0&1&0\\0&1&0&0\end{bmatrix}\qquad Q_4=\begin{bmatrix}0&0&0&1\\0&1&0&0\\1&0&0&0\\0&0&1&0\end{bmatrix}$$

$$Q_5=\begin{bmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\qquad Q_6=\begin{bmatrix}0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&0&1\end{bmatrix}$$

$$Q_7=\begin{bmatrix}0&0&1&0\\0&1&0&0\\0&0&0&1\\1&0&0&0\end{bmatrix}\qquad Q_8=\begin{bmatrix}0&1&0&0\\0&0&0&1\\0&0&1&0\\1&0&0&0\end{bmatrix}$$

显然，X 中任何一个元素都可以用 Q_1，Q_2，…，Q_8 来线性表示，它们能否构成 X 的一组基，关键在于它们是否线性无关。

容易看出

$$Q_1+Q_4+Q_5+Q_8-Q_2-Q_3-Q_6-Q_7=0$$

所以 Q_1，Q_2，…，Q_8 这 8 个基本方阵是线性相关的，即至少存在一个 Q_j，可以通过其他 7 个基本方阵的线性组合得到，这 8 个基本方阵的地位是等同的，故可不妨设 $j=8$。下面验证 Q_1，Q_2，…，Q_7 是否线性相关。

令 $\sum_{i=1}^{7} r_i Q_i = 0$，即

$$\begin{bmatrix}r_1+r_2 & r_6 & r_5+r_7 & r_3+r_4\\ r_3+r_5 & r_4+r_7 & r_1+r_6 & r_2\\ r_4+r_6 & r_2+r_5 & r_3 & r_1+r_7\\ r_7 & r_1+r_3 & r_2+r_4 & r_5+r_6\end{bmatrix}=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

等号两边对应元素相比较，得 $r_1=r_2=\cdots=r_7=0$，所以 Q_1，Q_2，…，Q_7 是线性无关的。Q_1，Q_2，…，Q_7 是空间的一组基，X 中任何元素都可以由 Q_1，Q_2，…，Q_7 的线性组合生成。可以这样认为：{ Q_1，Q_2，…，Q_8}是 X 的生成集，但不是最小生成集，而{ Q_1，Q_2，…，Q_7}是 X 的最小生成集。

现在，我们回到 Albrecht Dürer 铸造的铜币，以 Q_1，Q_2，…，Q_7 的线性组合表示铜币上的魔方，$D= d_1Q_1+d_2Q_2+\cdots+d_7Q_7$，即解方程组

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 5 & 15 & 14 & 1 \end{bmatrix} = \begin{bmatrix} d_1+d_2 & d_6 & d_5+d_7 & d_3+d_4 \\ d_3+d_5 & d_4+d_7 & d_1+d_6 & d_2 \\ d_4+d_6 & d_2+d_5 & d_3 & d_1+d_7 \\ d_7 & d_1+d_3 & d_2+d_4 & d_5+d_6 \end{bmatrix}$$

解得 $D=8Q_1+8Q_2+7Q_3+6Q_4+3Q_5+3Q_6+5Q_7$。

3.2.3 X 空间的子空间和 X 空间的扩展

改变对 Dürer 魔方数字和的要求，我们可以利用线性子空间的定义，构造 X 的子空间或者构造新的空间包含 X 空间。这里，我们规定仅包含 0 方的向量空间维数为零。

（1）要求数字方的所有数都相等。这是集合 $G=\{rE, r\in \mathrm{H}\}$，$G$ 是以 $\beta_G=\{E\}$ 为基的一维向量空间，是 X 的一维子空间。

（2）要求列和、行和及每条主、副对角线上数字和都相等，得到 5 维泛对角方的向量空间 B。例如：

$$P=\begin{bmatrix} 17 & 2 & 11 & 16 \\ 16 & 11 & 22 & -3 \\ 12 & 7 & 6 & 21 \\ 1 & 26 & 7 & 12 \end{bmatrix} \qquad H=N=R=C=S=46$$

其中 H 为主对角线和，N 为副对角线和。

它的基 B_B 为

$$P_1=\begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad P_2=\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad P_3=\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$P_4=\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad P_5=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

（3）要求行和、列和及两条对角线上的元素和相等，得到 8 维向量空间 Q，基向量 $Q_B=\{Q_1, Q_2, \cdots, Q_7, N_0\}$，其中 $Q_1, Q_2, \cdots, Q_7$ 是 X 的基，

$$N_0=\begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix}$$

例如，

$$T=\begin{bmatrix}6&7&9&8\\12&6&5&7\\5&10&9&6\\7&7&7&9\end{bmatrix}\qquad R=C=D=30$$

（4）仅要求行和与列和相等，可以得到 10 维向量空间 ψ，它的基 $\psi_B=\{Q_1, Q_2, \cdots, Q_7, N_1, N_2, N_3\}$，其中 $Q_1, Q_2, \cdots, Q_7$ 是 X 的基，而

$$N_1=\begin{bmatrix}0&0&0&0\\1&0&0&-1\\-1&0&0&1\\0&0&0&0\end{bmatrix}\qquad N_2=\begin{bmatrix}0&1&0&-1\\1&0&-1&0\\-1&0&0&1\\0&-1&1&0\end{bmatrix}$$

$$N_3=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}$$

（5）如果我们对数字没任何要求，那么所有的 4×4 数字方组成的向量空间 M，它的维数是 16，基向量 M_B 中的元素应是标准基（即仅有一个元素为 1，其余元素均为 0 的方阵）。

Botsch（1976 年）证明了可以构造大量的 X 的子空间或 X 的扩张空间。对于 1 与 16 之间的每一个数 K，都存在 K 维的 4×4 方的向量空间，其中的每一方阵都具有某些特定的性质。

由上可知，有下式成立

向量空间：$\{0\}\subseteq G\subseteq B\subseteq X\subseteq Q\subseteq \Psi\subseteq M$

维　　数：0　　1　5　7　8　10　16

3.3　密码的设计、解码与破译

密码的设计和使用至少可以追溯到4000多年前的古埃及、古巴比伦、古罗马和古希腊，历史极为久远。古代隐藏信息的方法主要有两大类：其一为隐藏信息载体，采用隐写术等；其二为变换信息载体，使之无法为一般人所理解。本节只涉及后者，介绍一些采用数学工具对信息加密、解密的方法。

在密码学中，信息代码被称为密码，加密前的信息被称为明文，经加密后不为常人所理解的用密码表示的信息被称为密文（ciphertext），将明文转变成密文的过程被称为加密（enciphering），其逆过程则被称为解密（deciphering），而用于加密、解密的方法或算法则被称为密码体制（crytosystem）。

记全体明文组成的集合为 U，全体密文组成的集合为 V，称 U 为明文空间，V 为密文空间。加密常利用某一被称为密钥的东西来实现，它通常取自一个被称为密钥空间的含有若干参数的集合 K。按数学的观点来看，加密与解密均可被看成是一种变换（或称映射）：取一 $k \in K$，$\forall u \in U$，令 $u \xrightarrow{k} v$ ÎV，v 为明文 u 在密钥 K 下的密文，而解码则要用到 K 的逆变换 K^1，$v \xrightarrow{k^{-1}} u$。由此可见，密码体系虽然可以千姿百态，但其关键还在于密钥的选取。

随着计算机与网络技术的迅猛发展，大量各具特色的密码体系不断涌现。离散数学、数论、计算复杂性、混沌……许多相当高深的数学知识都被用上，逐步形成了（并仍在迅速发展的）具有广泛应用面的现代密码学。

本节不准备涉及分组加密算法、身份证与消息认证、数字签名、椭圆曲线密码、量子密码、混沌密码、序列密码等具有某些独特功能的密码新体制。在这里我们只对古典密码、希尔密码作一个简要介绍，它们仅用到较为简单的线性代数知识。对需要用到其他数学知识的现代密码有兴趣的读者可以参阅密码学方面的专门书籍。

早期密码大体可分三类：代替法密码、移位密码和代数密码。

3.3.1 代替法密码

代替法密码采用另一个字母表中的字母来代替明文中的字母，明文字母与密文字母保持一一对应关系，但采用的符号改变了。加密时，把明文换成密文，即把明文中的字母用密文字母表中对应位置上的字母取代。解密时，则把密文换成明文，即把密文中的字母用明文字母表中对应位置上的字母代回，解密过程是加密过程的逆过程。在代替法加密过程中，明文字母表、密文字母表及两者间的对应关系即为代替法密钥，密钥既可以采用标准字母表，也可以任意建立。例如，我们可采用以下的明文字母表和密文字母表：

明文字母表　ABCDEFGHIJKLMNOPQRSTUVWXYZ

密文字母表　KLMNOPQRSTUVWXYZABCDEFGHIJ

密钥还经常用一密钥字或密钥短语生成混淆字母表。密钥字或密钥短语可以存放在识别码、通行字或密钥的秘密表格中。混合一个字母表，常见的有两种方法，这两种方法都采用了一个密钥字或一个密钥短语。

方法一

（1）选择一个密钥字或密钥短语，例如：CONSTRUCT

（2）去掉其中重复的字母，得：CONSTRU

（3）在修改后的密钥字后面接上从标准字母表中去掉密钥中的已有字母后剩下

的字母，得：

明文字母表　ABCDEFGHIJKLMNOPQRSTUVWXYZ

密文字母表　CONSTRUABDEFGHIJKLMPQVWXYZ

在设计密钥时，也可在明文字母表中选择一个特定字母，然后从该特定字母开始写密钥字并将密钥字隐藏于其中。例如，对于上例，选取特定字母 K，则可得：

明文字母表　ABCDEFGHIJKLMNOPQRSTUVWXYZ

密文字母表　KLMPQVWXYZCONSTRUABDEFGHIJ

方法二

（1）选择一个密钥字或密钥短语，例如：CONSTRUCT

（2）去掉其中重复的字母，得：CONSTRU

（3）这些字母构成矩阵的第一行，矩阵的后续各行由标准字母表中去掉密钥字的字母后剩下的字母构成：

$$\begin{bmatrix} C & O & N & S & T & R & U \\ A & B & D & E & F & G & H \\ I & J & K & L & M & P & Q \\ V & W & X & Y & Z & & \end{bmatrix}$$

（4）把所得矩阵中的字母按列的顺序选出，得

CAIVOBJWNDKXSELYTFMZRGPUHQ

按照此方法产生的字母表称为混淆字母表。

在代替法加密中，除了使用混淆字母表外，还可以使用混淆数。混淆数由以下方法产生：

（1）选一密钥字或密钥短语，例如：CONSTRUCT。

（2）按照这些字母在标准字母表中出现的相对顺序给它们编号，对序列中重复的字母则自左向右编号，得

$$\begin{matrix} C & O & N & S & T & R & U & C & T \\ 1 & 4 & 3 & 6 & 7 & 5 & 9 & 2 & 8 \end{matrix}$$

（3）自左向右选出这些数字，得到一混淆数字组：143675928，混淆字母表由从小到大的顺序取矩阵中相应列得出，先取第一列……再取第 8 列，依次得出秘文字母表。

为增加保密性，在使用代替法时还可利用一些其他技巧，如单字母表对多字母表、单字母对多字母、多重代替等，这里就不再一一细说了。

3.3.2 移位密码

移位密码采用移位法进行加密，明文中的字母重新排列，本身不变，只是位置改变了。

现在所知的最为古老的加密方法——天书就是移位法的一种。早在 4000 多年前，古希腊人就用一种名叫“天书”的器械来加密消息。该密码器械是用一条窄长的草纸缠绕在一个直径确定的圆筒上，明文逐行横写在纸带上，当取下纸带时，字母的次序就被打乱了，消息得以隐蔽。收方阅读消息时，要将纸带重新绕在直径与原来相同的圆筒上，才能看到正确的消息，在这里圆筒的直径起到了密钥的作用。

另一种移位法采用将字母表中的字母平移若干位的方法来构造密文字母表，传说这类方法是由古罗马皇帝恺撒最早使用的，故这种密文字母表被称为恺撒字母表。例如，如用将字母表向右平移 3 位的方法来构造密文字母表，可得：

明文字母表：ABCDEFGHIJKLMNOPQRSTUVWXYZ

密文字母表：DEFGHIJKLMNOPQRSTUVWXYZABC

使用这一密文字母表加密，THANK YOU 被加密成 WKDQN BRX，起到了一定的保密作用。

以上两种移位极易被人破译，为打破字母表中原有的顺序还可采用所谓路线加密法，即把明文字母表按某种既定的顺序安排在一矩阵中，然后用另一种顺序选出矩阵中的字母来产生密文表。

例如，对明文 THE HISTORY OF ZJU IS MORE THAN ONE HUNDRED YEARS。以 7 列矩阵表示如下：

THEHIST
ORYOFZJ
UISMORE
THANONE
HUNDRED
YEARS

再按事先约定的方式选出密文。例如，如按列选出，得到密文

TOUTHYHRIHUEEYSANAHOMNDRIFOORSSZRNETJEED

使用不同的顺序进行编写和选择，可以得到各种不同的路线加密体制。对于同一明文消息矩阵，采用不同的抄写，得到的密文也是不同的。如果对上例明文消息矩阵从左上角开始沿对角线抄写，得到的密文是

TOHURETIYHHHSOIYUAMFSENNOZTADORJRRNESEED

在使用上述方法时矩阵的规模必须事先约定，它增加了对明文的保护程度。当明文超过规定矩阵的大小时，可以另加一矩阵。当需要加密的字母数小于矩阵大小时，可以在矩阵中留空位或以无用的字母来填满矩阵。

移位法也可和代替法结合使用，并使用约定的单词或短语作密钥，以进一步加强保密性，这就是钥控列序加密法。例如，用密钥字 CONSTRUCT 对明文

MATHEMATICAL　MODELING　IS　USEFUL

加密：

1	4	3	6	7	5	9	2	8
C	*O*	*N*	*S*	*T*	*R*	*U*	*C*	*T*
M	*A*	*T*	*H*	*E*	*M*	*A*	*T*	*I*
C	*A*	*L*	*M*	*O*	*D*	*E*	*L*	*I*
N	*G*	*I*	*S*	*U*	*S*	*E*	*F*	*U*
L								

按混淆数的顺序选出各列，得到密文

MCNLTLFTLIAAGMDSHMSEOSIIUAEE

矩阵的最后一行可以用无用的字母填满，但若不加字母，则保密程度可以有所提高。移位法的使用可重复多次，只进行一次移位加密的称为一次移位法，经多次移位的则称为多次移位法。

3.3.3　代替法与移位法密码的破译

对窃听到的密文进行分析时，穷举法和统计法是最基本的破译方法，其他特殊的方法大多是这两种方法的综合和改进。

穷举法就是对所有可能的密钥或明文进行逐一试探，直至试探到“正确”的为止。此方法需要事先知道密码体制或加密算法（但不知道密钥或加密的具体办法）。破译时需将猜测到的明文和选定的密钥输入给算法，产生密文，再将该密文与窃听来的密文比较。如果相同，则认为该密钥就是所要求的，否则继续试探，直至破译。以英文字母为例，当已知对方在采用代替法加密时，如果使用穷举字母表来破译，那么对于最简单的一种使用单字母表-单字母-单元代替法加密的密码，字母表的可能情况有 26! 种，可见，单纯地使用穷举法，在实际应用中几乎是行不通的，只能与其他方法结合使用。

统计法是根据统计资料进行猜测的。在一段足够长且非特别专门化的文章中，字母的使用频率是比较稳定的。在某些技术性或专门化文章中的字母使用频率可能有微小变化。

在上述两种加密方法中，字母表中的字母是一一对应的，因此，在截获的密文中各字母出现的概率提供了重要的密钥信息。根据权威资料报道，可以将 26 个英文字母按其出现的频率大小较合理地分为五组：

Ⅰ. t,a,o,i,n,s,h,r;

Ⅱ. e;

Ⅲ. d,l;

Ⅳ. c,u,m,w,f,g,y,p,b;

Ⅴ. v,k,j,x,q,z;

不仅单个字母以相当稳定的频率出现，相邻字母对和三字母对同样如此。按频率大小将双字母排列如下：

th,he,in,er,an,re,ed,on,es,st,en,at,to,nt,

ha,nd,ou,ea,ng,as,or,ti,is,er,it,ar,te,se,hi,of

使用最多的三字母按频率大小排列如下：

the,ing,and,her,ere,ent,tha,nth,was,eth,for,dth

统计的章节越长，统计结果就越可靠。对于只有几个单词的密文，统计是无意义的。

下面介绍一下统计观察的三个结果：

（1）单词 the 在这些统计中有重要的作用；

（2）以 e，s，d，t 为结尾字母的英语单词超过了一半；

（3）以 t，a，s，w 为起始字母的英语单词约为一半。

对于（1），如果将 the 从明文中删除，那么 t 的频率将要降到第二组中其他字母之后，而 h 将降到第三组中，并且 th 和 he 就不再是最众多的字母了。

以上对英语统计的讨论是在仅涉及 26 个字母的假设条件下进行的。实际上消息的构成还包括间隔、标点、数字等字符。总之，破译密码并不是件很容易的事。

3.3.4 代数密码（希尔密码）

代替密码与移位密码的一个致命弱点是明文字符和密文字符有相同的使用频率，破译者可根据统计出来的字符频率找到规律，进而找出破译的突破口。要克服这一缺陷，提高保密程度就必须改变字符间的一一对应。1929 年，希尔利用线性代数中的矩阵乘积运算，打破了字符间的对应关系，设计了一种被称为希尔密码的代数密码。为了便于计算，希尔首先将字符变换成数，例如，对英文字母，我们可以作如下变换：

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
T	U	V	W	X	Y	Z												
20	21	22	23	24	25	0												

希尔密码的基本思想很简单，将密文分成 n 个一组，用对应的数字代替，就变成了一个个 n 维向量。如果取定一个 n 阶的非奇异矩阵 A（此矩阵为主要密钥），用 A 去乘每一向量，即可起到加密的效果，解密也不麻烦，将密文也分成 n 个一组，同样变换成 n 维向量，只需用 A^{-1} 去乘这些向量，即可将它们变回原先的明文（因为 $AA^{-1}=I$）。

在具体实施时，读者很快会发现一些困难：

（1）为了使数字与字符间可以互换，必须使用取自 0~25 之间的整数；

（2）由线性代数知识，$A^{-1}=\dfrac{A^*}{\det(A)}$，其中 A^* 为 A 的伴随矩阵，这说明在求 A 的逆矩阵时可能会出现分数。不解决这些困难，上述想法仍然无法实现。解决的办法是引进同余运算，并用乘法来代替除法（如同线性代数中用逆矩阵代替矩阵除法一样）。

让我们从最简单的情况做起，令 n=1，用数 a 去乘 0~25 中的数，以 26 为模取同余，并要求存在 $a^{-1}\in\{0,\cdots,25\}$，使得 $\forall p\in\{0,\cdots,25\}$，有 $a^{-1}ap\equiv p\ (\mathrm{mod}\ 26)$，或要求存在 a^{-1}，使得 $a^{-1}a=aa^{-1}\equiv 1\ (\mathrm{mod}\ 26)$。经简单的分析即可发现，并非所有 0~25 中的数都可用作这里的 a，事实上我们可证明下面的定理：

定理 3.1　$a\in\{0,1,\cdots,25\}$,若 $\exists a^{-1}\in\{0,1,\cdots,25\}$，使得 $aa^{-1}=a^{-1}a\equiv 1\ (\mathrm{mod}\ 26)$，则必有 $\gcd\{a,26\}=1$，其中 $\gcd\{a,26\}$ 为 a 与 26 的最大公因子。

证明　任取 $p\in\{0,\cdots,25\}$，令 $ap=26k+q$，于是

$$a^{-1}ap\equiv a^{-1}q\ (\mathrm{mod}\ 26)$$

又 $a^{-1}q=a^{-1}(ap-26k)=p-26a^{-1}k$

故 $(a^{-1}a-1)p=-26a^{-1}k$，由 p 的任意性可知必有

$$a^{-1}a\equiv 1\ (\mathrm{mod}\ 26)$$

即　$\exists k^*,a^{-1}a=26k^*+1$

上式又说明必有 $\gcd\{a,26\}=1$，不然它将整除 1，而这是不可能的。

此外，我们还不难证明这样的 a^{-1} 还是由 a 唯一确定的。

事实上，设有

$$a_1^{-1}a=26k_1+1\quad 和\quad a_2^{-1}a=26k_2+1\ \ (a_1^{-1},a_2^{-1}\in\{0,1,\cdots,26\}),$$

则 $(a_1^{-1}-a_2^{-1})a=26(k_1-k_2)$，故必有 $k_1=k_2$（也因为 $\gcd\{a,26\}=1$），即 $a_1^{-1}=a_2^{-1}$。

由定理 3.1，0~26 中除 13 以外的奇数均可取作这里的 a，下表为经计算求得的逆元素：

a	1	3	5	7	9	11	15	17	19	21	23	25
a^{-1}	1	9	21	15	3	19	7	23	11	5	17	25

例 3.3 取 $a=3$，用希尔密码体系加密语句 THANK YOU。

步骤 1 将 THANK YOU 转换成（20，8，1，14，11，25，15，21）。

步骤 2 每一分量乘以 3 并关于 26 取余得（8，24，3，16，7，23，19，11）

密文为 HXCPG WSK。

现在，我们已不难将方法推广到 n 为一般整数的情况，只需在乘法运算中结合应用取余，求逆矩阵时用逆元素乘来代替除法即可。

例 3.4 取 $A=\begin{bmatrix}1 & 2\\0 & 3\end{bmatrix}$，则 $A^{-1}=\begin{bmatrix}1 & 8\\0 & 9\end{bmatrix}$，用 A 加密 THANK YOU，再用 A^{-1} 对密文解密。

解 （1）希尔密码加密

用相应数字代替字母，划分为两个一组并表示为向量

$$\begin{bmatrix}8\\20\end{bmatrix}\quad\begin{bmatrix}1\\14\end{bmatrix}\quad\begin{bmatrix}11\\25\end{bmatrix}\quad\begin{bmatrix}15\\21\end{bmatrix},$$

用矩阵 A 左乘各向量加密（关于 26 取余）得

$$\begin{bmatrix}10\\24\end{bmatrix}\quad\begin{bmatrix}3\\16\end{bmatrix}\quad\begin{bmatrix}9\\23\end{bmatrix}\quad\begin{bmatrix}5\\11\end{bmatrix}$$

得到密文

JXCPI WEK

（2）希尔密码解密

用 A^{-1} 左乘得到的向量，即可还原为原来的向量，读者可自行检验之。

希尔密码是以矩阵法为基础的，明文与密文的对应由 n 阶矩阵 A 确定。矩阵 A 的阶数是事先约定的，与明文分组时每组字母的字母数量 n 相同，如果明文所含字数与 n 不匹配，则最后几个分量可任意补足。

（3）A^{-1} 的求法

方法一 利用公式 $A^{-1}=\dfrac{A^*}{\det(A)}$，例如，若取 $A=\begin{bmatrix}1 & 2\\0 & 3\end{bmatrix}$，则 $\det(A)=3$，$\dfrac{1}{\det(A)}=9$。

$$A^{-1}=9\begin{bmatrix}3 & -2\\0 & 1\end{bmatrix}(\bmod 26)\text{，即}\quad A^{-1}=\begin{bmatrix}1 & 8\\0 & 9\end{bmatrix}\text{。}$$

方法二　利用行初等变换的高斯消去法。将矩阵$(A\ ,I)$中的矩阵 A 消为 I，则原先的 I 即被消成了 A^{-1}，如求 $A=\begin{bmatrix}1 & 2\\0 & 3\end{bmatrix}$的逆矩阵可用如下方法：

写出$\left(\begin{bmatrix}1 & 2\\0 & 3\end{bmatrix},\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$，用 9 乘第二行化第二行的 3 为 1：

$$\left(\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix},\begin{bmatrix}1 & 0\\0 & 9\end{bmatrix}\right)$$

再从第一行减去第二行的 2 倍，并取同余得

$$\left(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},\begin{bmatrix}1 & 8\\0 & 9\end{bmatrix}\right)$$

此时，括号内左边的矩阵为单位阵，而右边的矩阵即为要求的 A^{-1}

希尔密码系统的解密必须知道该体系是用矩阵加密的并依赖于以下几把钥匙（Key）：

Key1　矩阵 A 的阶数 n，即明文是按几个字母划分的。

Key2　变换矩阵 A，只有知道了 A 才可能推算出 A^{-1}。

Key3　明文和密文由字母表转换成 n 维向量所对应的非负整数表(上面为方便起见，我们采用了字母的自然顺序，事实上，在采用希尔密码时，我们仍可兼用代替法，以便增加破译的难度)。

3.3.5　希尔密码的破译

希尔密码体系为破译者设置了多道关口，加大了破译难度。破译和解密是两个不同的概念，虽然两者同样是希望对密文加以处理得到明文的内容，但是它们有一个最大的不同——破译密码时，解密必需用到的钥匙未能取得，破译密码的一方需要依据密文的长度，文字的本身特征，以及行文习惯等各方面的信息进行破译。破译密码虽然需要技术，但更加重要的是“猜测”的艺术。“猜测”的成功与否直接决定着破译的结果。

破译希尔密码的关键是猜测文字转换成 n 维向量所对应的字母表，最重要的是获得加密矩阵 A。

由线性代数的知识可以知道，矩阵完全由一组基的变换决定，对于 $n\times n$ 矩阵 A，只要猜出密文中 n 个线性无关的向量 $q_i=Ap_i$ $(i=1,\ 2,\ \cdots,\ n)$对应的明文 p_i $(i=1,\ 2,\ \cdots,n)$，即可确定 A，将密码破译。

在实际计算中，可以利用以下方法：

令

$$P=(p_1, L, p_n), Q^{-1}=A(p_1, L, p_n)=AP^{-1}$$

则

$$Q=PA^{-1}, \quad P=Q(A^{-1})^{-1}$$

取矩阵$[Q \mid P]$，经过一系列初等行变换，将由密文决定的 n 维矩阵 Q 化为单位阵 $I\left(n\times n\right)$ 的时候，由明文决定的矩阵 P 自动化为$(A^{-1})^{-1}$：

$$[Q,P]=[Q,Q(A^{-1})^{-1}],$$

初等行变换

$$[Q^{-1}Q,Q^{-1}Q(A^{-1})^{-1}]=[I,(A^{-1})^{-1}]$$

例如有密文如下：

goqbxcbuglosnfal;

根据英文的行文习惯以及获取密码的途径和背景，猜测是两个字母为一组的希尔密码，前四个明文字母是 dear；则前两组明文字母 de 和 ar 对应的二维向量是

$$P_1=(4,5)^{-1}, \quad P_2=(1,18)^{-1}。$$

按同一对应整数表，密文中对应这两组的二维向量是：

$$Q_1=AP_1=(7,15)^{-1}, \quad Q_2=AP_2=(17,2)^{-1}。$$

利用这一逆矩阵，可对截获的密文进行解密，破译出的电文是 Dear Mac God forbid。

这只是对最简单的情况进行的举例，如果加密矩阵的阶数大于 2，需要的密文应该有较长长度，所需进行的计算量也是很大的，但在猜测到 P_1, L, P_n 后求解加密矩阵的计算量却并不算大，是与 n^2 同阶的。

本节中介绍的希尔密码体制是以矩阵法为基础的，实际上，希尔首先提出的是联立方程法，而后经简化提出了利用矩阵进行运算的希尔密码，由线性代数知识，这一点不难理解。

希尔密码加密、解密、破译的过程中有两个要素非常重要：第一个要素是字母与 n 维向量进行转换所依据的非负整数表，本节中所举的是最自然的情况；当然如果依据其他的整数表也是完全可以进行的，其情况将会更复杂一些，破译的难度就会增大。

第二个要素是加密矩阵，如何定义、求解这个矩阵对于密码的加密和破译更加关键。如前所述，加密时应选择对应行列式的值与 26 没有公因子的矩阵，使用时可

以任意选择。

习　题　三

1.试证明：在小船只能乘坐 2 人时，4 对阿拉伯夫妇是无法过河的。

2.假如河中有一个小岛，4 对夫妇即可过河，请给出一个可行方案。

3.作出一个 7 阶魔方。

4.请给出一个不同于$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$的 2 阶希尔矩阵 A。

（1）利用你给出的矩阵 A 加密短语 Thank you。

（2） 求出相应的解密矩阵 A^{-1}，你不妨可验证一下，利用它解密，是否可还原出 Thank you。

第4章　微分方程建模

4.1　微分方程的一般理论

微分方程是研究函数变化规律的有力工具，有着广泛的实际应用。针对所研究的对象建立微分方程模型是解决问题的第一步，实际中只有求出微分方程的解才能对所研究的问题进行解释说明。一般说来，求微分方程的解析解是困难的，大多数的微分方程需要用数值方法来求解，因此首先需要研究微分方程的解的存在唯一性和稳定性问题。

4.1.1　微分方程的一般形式

一阶微分方程

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=f(t,x)\\ x(t_0)=x_0\end{cases} \tag{4.1}$$

其中 $f(t,x)$ 是 t 和 x 的已知函数，$x(t_0)=x_0$ 为初始条件，又称定解条件。

一阶微分方程组

$$\begin{cases}\dfrac{\mathrm{d}x_i}{\mathrm{d}t}=f_i(t,x_1,x_2,\cdots,x_n)\quad(i=1,2,\cdots,n)\\ x_i(t_0)=x_i^{(0)}\quad(i=1,2,\cdots,n)\end{cases} \tag{4.2}$$

又称为一阶正规方程组。若引入向量

$$x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}},\ x_0=(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)})^{\mathrm{T}},$$

$$f=(f_1,f_2,\cdots,f_n)^{\mathrm{T}},\ \frac{\mathrm{d}x}{\mathrm{d}t}=\left(\frac{\mathrm{d}x_1}{\mathrm{d}t},\frac{\mathrm{d}x_2}{\mathrm{d}t},\cdots,\frac{\mathrm{d}x_n}{\mathrm{d}t}\right)^{\mathrm{T}},$$

则方程组（4.2）可以写为简单的形式

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=f(t,x)\\ x(t_0)=x_0\end{cases} \tag{4.3}$$

即与方程（4.1）的形式相同，当 $t\in[t_{k-1},\ t_k]$ 时为方程（4.1）。

对于任一高阶的微分方程

$$\frac{d^n x}{dt^n}=f\left(t;x,\frac{dx}{dt},\cdots,\frac{d^{n-1}x}{dt^{n-1}}\right),$$

若记 $\frac{d^i x}{dt^i}=y_i\left(i=0,1,\cdots,n\right)$，则方程为

$$\frac{dy_{n-1}}{dt}=f\left(t;y_0,y_1,\cdots,y_{n-1}\right),$$

即可化为一阶方程组的形式。因此，下面主要对正规方程组（4.3）进行讨论。

4.1.2　微分方程解的存在唯一性

正规方程组（4.3）的解在什么条件下存在，且唯一呢？有下面的定理。

定理 4.1　（Cauchy-Peano）若函数 $f(t,x)$ 在 $R:|t-t_0|\le a,|x-x_0|\le b$ 上连续，则方程组（4.3）在 $|t-t_0|\le h$ 上有解 $x=\phi(t)$ 满足初值条件 $x_0=\phi(t_0)$，此处

$$h=\min\left(a,\frac{b}{M}\right),M=\max_{(t,x)\in R}|f(t,x)|$$

定理 4.2　若函数 $f(t,x)$ 在 $R:|t-t_0|\le a,|x-x_0|\le b$ 上连续，且满足利普希茨（Lipschitz）条件（即存在正常数 L 使得 $|f(t,x^{(1)})-f(t,x^{(2)})|\le L|x^{(1)}-x^{(2)}|$，其中 $(t,x^{(1)}),(t,x^{(2)})\in \mathbf{R}$），则方程组（4.3）满足初值条件 $x_0=\phi(t_0)$ 的解是唯一的。

4.1.3　微分方程的稳定性问题

在实际问题中，微分方程所描述的是物质系统的运动规律，在用微分方程来研究这个物理过程中，人们只能考虑影响该过程的主要因素，而不得不忽略一些次要因素，这些次要因素通常称为干扰因素。这些干扰因素在实际中可能瞬时地起作用，也可能持续地起作用。从数学上来看，前者会引起初值条件的变化，而后者则会引起微分方程本身的变化。在实际问题中，干扰因素是客观存在的，由此可见，对于它的影响程度的研究是必要的，即初值条件或微分方程的微小变化是否也只引起对应解的微小变化？这就是微分方程的稳定性问题。这里仍以方程组（4.3）为例讨论。

1.有限区间的稳定性

若 $f(t,x)$ 在某个有限的区域 $G\subset R^{n+1}$ 内连续，且对 x 满足利普希茨条件，$x=\psi(t)\left(a\le t\le b\right)$ 是方程组（4.3）的一个特解，则当 x_0 充分接近于 $\psi(t_0)\left(a\le t_0\le b\right)$ 时，方程组（4.3）在 $a\le t_0\le b$ 上满足初值条件 $x_0=x(t_0)$ 的解 $x=\phi$（t, t_0, x_0）有

$$\lim_{x_0\to\psi(t_0)}\phi(t,t_0,x_0)=\psi(t)\left(a\le t\le b\right)$$

即对任意给定的 $\varepsilon>0$，总存在相应的 $\delta(\varepsilon)>0$，当 $|x_0-\psi(t_0)|<\delta(\varepsilon)$ 时，对一切 $t\in[a,b]$，有

$$\left|\phi(t,t_0,x_0)-\psi(t)\right|<\varepsilon$$

此时称方程组（4.3）的解 $x=\psi(t)$ 在有限区间 $[a,b]$ 上是稳定的。

2.无限区间的稳定性

若 $x=\psi(t)$ $(t_0\le t)$ 是方程组（4.3）的一个特解，$x=\phi(t,t_0,x_0)$ $(t_0\le t)$ 是方程组（4.3）满足初值条件 $x_0=x(t_0)$ 的解。对任意给定的 $\varepsilon>0$，总存在相应的 $\delta(\varepsilon)>0$，当 $\left|x_0-\psi(t_0)\right|<\delta(\varepsilon)$ 时，对一切 $t\ge t_0$ 有 $\left|\phi(t,t_0,x_0)-\psi(t)\right|<\varepsilon$，则称方程组（4.3）的解 $x=\psi(t)$ 在无限区间 $[t_0,+\infty)$ 上是稳定的。

3.渐近稳定性

若方程组（4.3）的解 $x=\psi(t)$ 在无限区间 $[t_0,+\infty)$ 上是稳定的，且存在 $\delta_0>0$，当 $\left|x_0-\psi(t_0)\right|<\delta_0$ 时，有

$$\lim_{t\to\infty}\left[\varphi(t,t_0,x_0)-\psi(t)\right]=0$$

则称 $x=\psi(t)$ 是渐近稳定的，或称局部渐近稳定性。

若上述 $\delta_0=\infty$（或给定的一个有限常数），则相应的渐近稳定性称为全局渐近稳定性（或大范围渐近稳定性）。

4.经常扰动下的稳定性

对于方程组（4.3），考虑相应的方程组

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(t,x)+R(t,x) \tag{4.4}$$

这里的 $R(t,x)$ 称为扰动函数。

若对任意给定的 $\varepsilon>0$，总存在 $\delta(\varepsilon)>0$ 和 $\eta(\varepsilon)>0$，使得当 $\left|x_0-\psi(t_0)\right|<\delta(\varepsilon)$ 时有

$$\left|R(t,x)\right|<\eta(\varepsilon)$$

则方程组（4.4）有满足初值条件 $x_0=x(t_0)$ 的解 $x=\phi(t,t_0,x_0)(t_0\le t)$。且当 $t\ge t_0$ 时有

$$\left|\varphi(t,t_0,x_0)-\psi(t)\right|<\varepsilon$$

就说方程组（4.3）的特解 $x=\psi(t)$ 在经常扰动下是稳定的。

5.研究稳定性的方法

实际中，要研究方程组（4.3）的解 $x=\psi(t)$ 的稳定性问题，可以转化为研究方程的零解（平凡解）的稳定性问题。事实上，对于方程组（4.3）的任一特解 $x=\psi(t)$，只要令 $y=x-\psi(t)$，则

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{\mathrm{d}x}{\mathrm{d}t}-\frac{\mathrm{d}\psi(t)}{\mathrm{d}t}&=f(t,x)-f(t,\psi(t))\\&=f(t,y+\psi(t))-f(t,\psi(t))=g(t,y)\end{aligned}$$

显然有 $g(t,0)\equiv0$。所以方程组（4.3）变为

$$\frac{\mathrm{d}y}{\mathrm{d}t}=g(t,y) \tag{4.5}$$

于是方程组（4.3）的解 $x=\psi(t)$ 对应于方程组（4.5）为 $y=0$（平凡解）。因此，要研究方程组（4.3）的解 $x=\psi(t)$ 的稳定性问题可转化为研究方程组（4.5）的平凡解 $y=0$ 的稳定性问题。

若微分方程组的所有解都能简单地求出来，一个特解的稳定性问题并不难解决。可是，实际中此种情况太少了。因此，一般性的稳定性问题的研究是复杂的，通常情况下都是针对具体问题做相应的研究。

4.2 微分方程的平衡点及稳定性

4.2.1 微分方程的平衡点

设有微分方程组（4.3），对于 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in R^n, t\in[a,b], f(t,x)$ 在某个区域内连续，且满足解的存在唯一性条件。若存在某个常数 $x_0\in R^n$，使得 $f(t,x_0)=0$，则称点 x_0 为方程组（4.3）的平衡点（或奇点），且称 $x=x_0$ 为方程组的平凡解（或奇解）。

若对所有可能的初值条件，方程组（4.3）的解 $x=\psi(t)$ 都满足

$$\lim_{t\to\infty}\psi(t)=x_0$$

则称平衡点 x_0 是稳定的（渐近稳定）；否则是不稳定的。

实际中，判断平衡点的稳定性有两种方法：间接方法和直接方法。

间接方法　首先求出方程的解 $x=\psi(t)$，然后利用定义 $\lim\limits_{t\to\infty}\psi(t)=x_0$ 来判断。

直接方法　不用求方程的解直接研究其稳定性。

4.2.2 一阶微分方程的平衡点及稳定性

设有微分方程 $\dfrac{\mathrm{d}x}{\mathrm{d}t}=f(x)$，其相应的平衡点为代数方程 $f(x)=0$ 的实根 $x=x_0$。其稳定性可用间接方法来判断，下面说明直接方法。

首先，将函数在点 x_0 作一阶泰勒（Taylor）展开，即方程可以近似地表示为

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f'(x_0)(x-x_0)$$

显然，x_0 也是该方程的一个平衡点，其稳定性主要取决于 $f(x_0)$ 的符号，即有下面结论：

若 $f'(x_0)<0$，则平衡点 x_0 是稳定的；若 $f'(x_0)>0$，则平衡点 x_0 是不稳定的。

4.2.3 平面方程组的平衡点及稳定性

设平面方程组的一般形式为

$$\begin{cases}\dfrac{dx_1}{dt}=f(x_1,x_2)\\ \dfrac{dx_2}{dt}=g(x_1,x_2)\end{cases} \tag{4.6}$$

则称代数方程组

$$\begin{cases}f(x_1,x_2)=0\\ g(x_1,x_2)=0\end{cases}$$

的实根 $x_1=x_1^{(0)},x_2=x_2^{(0)}$ 为平面方程组（4.6）的平衡点，记为 $P_0\left(x_1^{(0)},x_2^{(0)}\right)$。若对所有可能的初值条件，方程的解 $x_1(t),x_2(t)$ 满足

$$\lim_{t\to\infty}x_1(t)=x_1^{(0)},\ \lim_{t\to\infty}x_2(t)=x_2^{(0)}$$

则称平衡点 $P_0\left(x_1^{(0)},x_2^{(0)}\right)$ 是稳定的；否则是不稳定的。亦可以用直接方法讨论。

将方程组（4.6）的右边的函数做一阶泰勒展开，即可表示为近似的线性方程组

$$\begin{cases}\dfrac{dx_1}{dt}=f_{x_1}\left(x_1^{(0)},x_2^{(0)}\right)\left(x_1-x_1^{(0)}\right)+f_{x_2}\left(x_1^{(0)},x_2^{(0)}\right)\left(x_2-x_2^{(0)}\right)\\ \dfrac{dx_2}{dt}=g_{x_1}\left(x_1^{(0)},x_2^{(0)}\right)\left(x_1-x_1^{(0)}\right)+g_{x_2}\left(x_1^{(0)},x_2^{(0)}\right)\left(x_2-x_2^{(0)}\right)\end{cases} \tag{4.7}$$

记系数矩阵为 $A=\begin{bmatrix}f_{x_1} & f_{x_2}\\ g_{x_1} & g_{x_2}\end{bmatrix}_{P_0}$，且假设其行列式 $|A|\neq 0$，则方程组（4.7）的特征方程为

$$|A-\lambda I|=0$$

即

$$\lambda^2+p\lambda+q=0$$

其中 $p=-\left(f_{x_1}+g_{x_2}\right)\big|_{P_0}$，$q=|A|$，$\lambda$ 为特征根。不妨设特征根分别为 λ_1,λ_2，即

$$\lambda_1,\lambda_2=\frac{1}{2}\left(-p\pm\sqrt{p^2-4q}\right)。$$

根据特征根 λ_1,λ_2 和系数 p，q 的取值情况可以确定平衡点 $P_0\left(x_1^{(0)},x_2^{(0)}\right)$ 的稳定性。

事实上，当 $p>0,q>0$ 时，平衡点是稳定的；当 $p<0$ 或 $q<0$ 时，平衡点是不稳定的。

对于一般微分方程的平衡点和稳定性问题可以类似地讨论。

4.3　基于一种复合模型的中国人口预测模型

4.3.1　研究背景

在我国落实科学发展观、构建和谐社会的进程中，必须实现人口与经济、社会、资源、环境协调发展和可持续发展，进一步合理控制人口数量，提高人口质量，改

善人口结构。因此，我们完全有必要建立一种较理想的人口总数预测模型，对我国人口总数进行中短期内准确的预测，为相关部门提供一定的数据，最终能够采取适当的措施控制人口发展，从而有利于我国经济的发展。试从我国历次人口普查数据和人口增长的特点出发，运用数学建模的思想方法建立我国人口预测的数学模型，并由此对我国人口发展的中长期趋势做出预测。

国外关于人口发展的模型主要有 Malthus 人口模型和 Logistic 人口模型，这两个模型有一个共同的特点，就是它们的积分曲线都为单调递增曲线，也就是它们只适用于人口持续增长的情形，不符合未来我国人口的发展状况。

国内关于人口预测的方法大致分为两类：一类是灰色 GM（1，1）预测模型，该模型能准确地预测短期内我国的人口总数，但对中长期预测的偏差较大且随着时间的推移愈来愈大。例如：2010 年我国人口总数实际值（普查数据）是 13.70537 亿，预测值是 13.4246 亿，绝对误差为 0.28077 亿，误差率高达 2.05%. 另一类是宋健等人建立的中长期人口发展方程的人口预测模型，分为人口发展方程的离散形式与连续形式。离散形式的人口模型非常贴近实际，特别是在人口增长趋势发生转变的情况下能够准确地反映变化。但是模型中需要确定大量的参数，需要比较多比较准确的数据，而这些数据的获取又有一定的难度，且数据也多少有些误差，所以导致预测方法较难实施。建立的人口发展方程的连续形式模型，由于求方程的准确解比较困难，所以应用于实际不方便。

综上所述，我们必须寻求新的方法，建立一种新的人口预测模型，能准确、简便地对我国人口总数进行中长期预测。为此，本文先在阻滞增长模型的基础上建立了一种复合模型，再通过对我国历次人口普查数据规律的研究，用一元线性回归法确定此复合模型中参数的值，得到一种能准确预测我国人口发展趋势的人口模型。

4.3.2　模型的条件假设

要构建出较理想的数学模型，必须提出下列合理的假设：

假设一　人口总数只与人口基数有关；

假设二　不考虑意外灾难等因素对人口变化的影响；

假设三　不考虑我国人口同境外人口之间的迁入与迁出问题；

假设四　假设短期内我国的计划生育政策不变；

假设五　为了使出生和死亡的人数在整个时间段内服从均匀分布，保证 $N(t)$ 为 t 的连续函数且关于 t 可微，我们假设人口总数 $N(t)$ 足够大；

假设六　假设上一次人口普查到下一次人口普查这一段时间内，人口年增长率取这一段时间内的平均值。

4.3.3 模型建立

我们先在 Logistic 模型的基础上建立一种预测人口总数的复合模型，然后在此基础上建立中国人口总数预测模型。

4.3.3.1 一种复合模型

1837 年荷兰生物学家 Verhulst 提出生物群体增长的逻辑律，可表示为方程

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t}=\alpha N(t)-\beta N(t)^2 \tag{4.8}$$

其中$\alpha>0,\beta>0$，都是常数，且α和β相比是个非常微小的数。

后来，人们又称非线性方程（4.8）为生物种群增长的 Logistic（逻辑斯蒂）模型，或阻滞增长模型。这一模型是在下列假设的基础上建立的：

（1）假设有一个环境条件允许的种群数量的最大值，这个数值称为环境容纳量或负荷量，通常用M表示。当种群数量达到M时，种群将不再增长，即$\frac{\mathrm{d}N(t)}{\mathrm{d}t}=0$。

（2）假设环境条件对种群的阻滞作用随着种群密度的增加而按比例增加。

根据阻滞增长模型，人口增长模型可表示为微分方程初值问题：

$$\begin{cases}\frac{\mathrm{d}N(t)}{\mathrm{d}t}=\left[\alpha-\beta N(t)\right]N(t)\\ N(t_0)=N_0\end{cases} \tag{4.9}$$

其中$P(t)=\alpha-\beta N(t)$是人口净增长率；α是人口自然增长率；β为生命系数；$N(t)$为t时刻的人口总量；$N(t_0)$为初始时刻t_0的人口总量。

（4.9）是一个伯努利（Bernoulli）方程的初值问题，我们可以求出它的解，但常常由于数据资料有限，导致参数 α、β不能确定。所以即使求出了它的解也不能用来预测人口总数。但是，如果人口净增长率和时间呈线性关系，即有人口净增长率模型：

$$P(t)=at+b,\ a,b\text{是参数},\ ab<0 \tag{4.10}$$

我们只要将模型（4.10）代入模型（4.9），就会得到一种反映人口总数变化趋势的模型：

$$\begin{cases}\frac{\mathrm{d}N(t)}{\mathrm{d}t}=(at+b)N(t)\\ N(t_0)=N_0\end{cases} \tag{4.11}$$

其中，a，b是参数，$ab<0$。

求解此微分方程初值问题得：

$$N(t)=N_0\cdot\exp\left[-\left(bt_0+\frac{a}{2}t_0^2\right)\right]\cdot\exp\left(bt+\frac{a}{2}t^2\right) \tag{4.12}$$

在微分方程（4.11）中，令 $\frac{\mathrm{d}N(t)}{\mathrm{d}t}=(at+b)N(t)=0$，由于 $N(t)>0$，所以 $t=-\frac{b}{a}$ 是函数 $N(t)$ 的唯一稳定点，因此是 $N(t)$ 的最值点。当 $a<0,b>0$ 时，$N(t)$ 在 $\left(0,\ -\frac{b}{a}\right]$ 内单调递增，在 $\left[-\frac{b}{a},+\infty\right)$ 内单调递减，所以 $N(t)$ 在 $t=-\frac{b}{a}$ 处取到最大值 $N\left(-\frac{b}{a}\right)$；当 $a>0,b<0$ 时，$N(t)$ 在 $\left(0,\ -\frac{b}{a}\right]$ 内单调递减，在 $\left[-\frac{b}{a},+\infty\right)$ 内单调递增，所以 $N(t)$ 在 $t=-\frac{b}{a}$ 处取到最小值 $N\left(-\frac{b}{a}\right)$。

由于模型（4.11）是由两个模型（4.9）和（4.10）复合而成的，所以称它为复合模型。

4.3.3.2　中国人口总量预测模型

人口发展过程线上任一点的切线斜率基本保持不变，即各时期人口发展速度较一致，近似直线状延伸时可采用一元线性回归法。这里将时间作为控制变量，将人口净增长率作为状态变量，确定它们之间的数学模型。

设 t 介于两个相邻的普查年份之间，根据文献记载的全国历次人口普查数据计算出人口年平均净增长率 $P(t)$，见表 4-1。

表 4-1　中国历次人口普查数据分析（人口总数单位：万人）

序号	年份	人口总量	增长量	净增长率	年平均净增长率
1	1953	60193.8	—	—	—
2	1964	72307	12113.2	0.201	0.017
3	1982	103188.3	30881.3	0.427	0.02
4	1990	116001.7	12813.4	0.124	0.015
5	2000	129533	13531.3	0.117	0.011
6	2010	137053.7	7520.7	0.0581	0.0057

由于我国前三次人口普查并非每十年一次，时间间隔不相等，这给我们研究问题带来了不便。但我们假设了两个相邻普查年份之间各年的人口增长率，取这一年段的平均值，所以我们选取下列五个点：

$(1963,0.017),(1974,0.02),(1985,0.015),(1996,0.011),(2007,0.0057)$，

构造一元线性回归

$$X=(1963\quad 1974\quad 1985\quad 1996\quad 2007)^{\mathrm{T}},$$

$$Y=(0.017\quad 0.02\quad 0.015\quad 0.011\quad 0.0057)^{\mathrm{T}}$$

利用 Mathcad 软件计算相关系数得

$$corr(X,Y)=-0.899$$

由于 $0.878<|corr(X,Y)|<0.959$，所以 X、Y 之间存在显著的线性关系，即年份 t 和人口净增长率 $P(t)$ 具有良好的线性关系，不妨设

$$P(t)\approx at+b$$

a、b 是参数。

运用 Mathcad 求得 $a=-2.873\times10^{-4}, b=0.584$，得到回归方程

$$r(t)=-2.873\times10^{-4}t+0.584$$

拟合的结果如图 4-1。

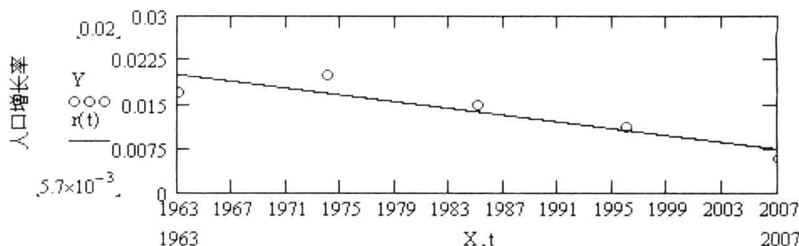


图 4-1 根据所选数据点拟合的人口增长率曲线

我们忽略这种误差，即令 $|P(t)-r(t)|=0$，于是得到我国人口净增长率模型：

$$P(t)=-2.873\times10^{-4}t+0.584 \tag{4.13}$$

我国从 1980 年开始实行计划生育政策。模型（4.13）刻画了两个阶段的人口净增长率，即不仅刻画了没有实行计划生育政策（1950—1979 年）时的人口净增长率，而且刻画了 1980 年以后的人口净增长率，也就是说反映我国人口净增长率的模型已包含了计划生育政策因素。所以将 $a=-2.873\times10^{-4}, b=0.584$ 代入复合模型（4.11）中，得到我国人口总数预测模型。

$$\begin{cases}\dfrac{\mathrm{d}N(t)}{\mathrm{d}t}=(0.584-2.873\times10^{-4}t)N(t)\\ N(t_0)=N_0\end{cases} \tag{4.14}$$

不妨选取初值条件 $N(2010)=137053.7$，连同 $a=-2.873\times10^{-4}, b=0.584$ 代入（4.12）式，得模型（4.14）的解：

$$N(t)=137053.7\exp(-0.584\times2010+1.4365\times10^{-4}\times2010^2+0.584t-1.4365\times10^{-4}t^2) \tag{4.15}$$

用 Mathcad 作出 （4.15）的曲线图（见图 4-2）。

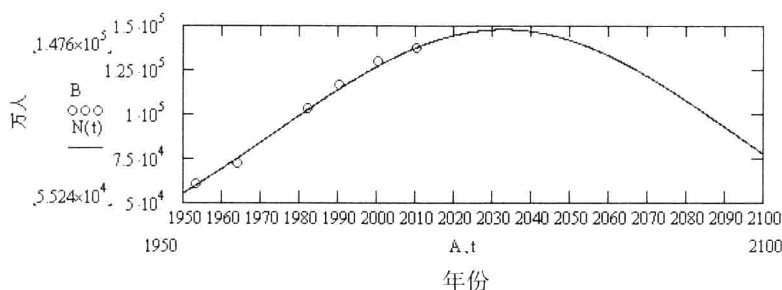


图 4-2　我国 1950—2100 年人口总数变化趋势图

由于 $a=-2.873\times10^{-4}<0, b=0.584>0$ ，所以根据前面的讨论， $N(t)$ 的最大值点是 $t=-\dfrac{b}{a}=2032.7184$ 。这表明我国人口将于 2033 年达到零增长。

图 4-2 显示，我国人口的增长并非 Logistic 人口模型所描述的那种简单的单调增加。在大概直到 2033 年的这一段时期内一直保持增加势头；人口最大值不超过 15 亿；大概从 2035 年开始进入负增长期。可见模型（4.14）是一种非单调的人口预测模型。

4.3.4 人口预测与模型验证

首先，我们用 Mathcad 在同一坐标系中分别作出六个点

（1953，60193.8），（1964，72307），（1982，103188.3），

（1990，116001.7），（2000，129533），（2010，137053.7）

和（4.15）式的图像，得到实际人口数据点与（4.15）式的拟合曲线图（见图 4-2）。

我们看到有 3 个点在曲线上，另外 3 个点在曲线的近旁。拟合结果非常好，说明模型（4.14）适合 1953—2010 年的我国人口总数增长状况。

其次，按照（4.15）式用数学软件 Mathcad 计算出我国 2005—2010 年的人口总数预测值，再与实际统计值进行比较（见表 4-2）。

表 4-2　我国 2005—2010 年的人口总数预测值与实际统计值的误差（单位：万人）

序号	年份	实际人口数	预测人口数	绝对误差	误差平方	误差率
1	2005	130756	132200	1444	2085136	0.0110
2	2006	131448	133200	1752	3069504	0.0133
3	2007	132129	134200	2071	4289041	0.0157
4	2008	132802	135200	2398	5750404	0.0181
5	2009	133474	136100	2626	6895876	0.0197
6	2010	137053.7	137100	46	2116	0.0003

根据表 4-2 中的误差平方和误差率分别计算出相对标准差 S 和平均误差

（*MAPE*）：

$$S=\sqrt{\frac{1}{6}\sum_{i=1}^{6}e_i^2}=1919\text{万人}；\quad MAPE=\frac{1}{6}\sum_{i=1}^{6}v_i=0.013$$

其中 e_i 表示绝对误差，v_i 表示误差率。

从表 4-2 可以看出，2010 年的绝对误差比其他年份的小得多，这是因为 2010 年的实际人口数是经第六次全国人口普查得到的，因而数据准确性高；其他五个年份的实际人口数是国家统计局抽样统计的结果，数据准确性不高。显然前五个年份的实际人口数的统计值偏小，因此使绝对误差增大，导致相对标准差 *S* 和平均误差率 *MAPE* 偏高。换言之，导致相对标准差和平均误差率较大的主要原因是国家统计局抽样统计的数据不准确。即使如此，模型（4.14）的精度还是比较高的，一方面，虽然 *S* 和 *MAPE* 的值较大，但也不是太大；另一方面，它毕竟是建立在我国历次人口普查数据基础之上的。

再次，用公式（4.15）预测我国 2011 年至 2038 年各年的人口总数，结果如表 4-3。

表 4-3　我国 2011—2038 年人口总数预测值（万人）

年份	总人口数	年份	总人口数	年份	总人口数	年份	总人口数
2011	137900	2018	143100	2025	146300	2032	147600
2012	138800	2019	143700	2026	146600	2033	147600
2013	139600	2020	144200	2027	146900	2034	147600
2014	140400	2021	144700	2028	147400	2035	147500
2015	141100	2022	145200	2029	147300	2036	147400
2016	141800	2023	145600	2030	147400	2037	147200
2017	142500	2024	146000	2031	147500	2038	147000

预测结果表明：我国在 2011 年到 2032 年这一段时期内人口一直将保持低增长态势；2033 年达到零增长，人口最大值为 147600 万；2035 年开始负增长。这些结论与近日国家统计局公布的“预计中国将在 2032 年前后达到人口零增长”非常吻合。

最后，国家人口发展战略研究报告指出：我国人口将于 2010 年、2020 年分别达到 13.6 亿人和 14.5 亿人，2033 年前后达到峰值 15 亿人左右。这与我们得到的在 2010 年人口为 13.71 亿人和 2020 年达到 14.42 亿人以及 2033 年的 14.76 亿人很接近。

以上四方面的原因足以说明我们所建模型（4.14）的合理性。

4.3.5　模型评价

我们在建立模型时是从人口增长率是时间的函数这一根本点出发，因而得到适合我国人口增长状况的人口总数预测模型，预测结果很接近实际值。但是模型还是有一定的误差。事实上，人口增长模型是由生育、死亡、疾病、自然灾害、特殊的人口生育政策、环境、社会以及经济等诸多因素影响和制约的结果，如此众多的因素不可能通过几个指标表达清楚，它们对人口增长的潜在而复杂的影响更是无法精确计算。这说明人口预测是一个相当复杂的问题。由于数据有限，未能充分考虑影响人口增长的因素，因此所建模型不全面，所得的结果与实际有一定的出入。

4.4　温室蔬菜杀虫剂合理使用方案

4.4.1　问题的提出

在温室中种植蔬菜的菜农们为了减少病虫害造成蔬菜的巨大损失，普遍使用杀虫剂。可是，不少菜农却发现使用农药的效果并不理想，使用不当不仅会浪费大量的人力和财力，还可能导致蔬菜中的农药残留量超标，不能食用。这种现象产生的主要原因之一就是不讲究科学用药，施药时没有抓准病虫害的有效防治期，没有控制好用药量和用药次数。因此，本书从经济学的角度出发，用定量方法讨论：在使用农药时如何抓准病虫害的有效防治期，控制好用药量和用药次数才能达到成本低、效果好的目的。希望能找到一种经济有效的杀虫剂使用方案，按此种方案使用农药，可以在防止蔬菜受其破坏的前提下有效地控制住病虫害，以达到提高经济效益的目的。

在一块占地面积为 S 公顷的温室内种植了某种蔬菜，该蔬菜的生长期为 T 天，按照蔬菜农药残留量的要求，在蔬菜生长期内每公顷每次农药施用量不得超过 M 千克。假设在蔬菜生长期初期，发现蔬菜上害虫的平均密度为每公顷 X_0 只，平均每只害虫每天大约吞食蔬菜 Q 千克，害虫每天的自然繁殖率约为 β。

我们希望知道的是：对该温室中的蔬菜应如何使用农药，即使用农药几次和什么时候使用以及每次使用多少千克农药，才能尽量降低害虫对蔬菜的破坏，并减少人工费、设备使用费等的投入?

4.4.2　问题分析

由于害虫不断地繁殖增长，因此需要进行定期杀虫。杀虫次数过于频繁、每次用药剂量过大，会导致药费与劳务费的增加；杀虫次数过少，害虫数量会猛增，导致大量蔬菜遭受破坏，使蔬菜损失费增大。因此，制定用药的最优策略，就是确定

合理的用药次数和每次的用药剂量，使总损失费用（蔬菜损失费、杀虫药费和人工费）最小。

4.4.3 模型的条件与假设

1.问题假设

要构建出较理想的数学模型，必须提出下列合理的假设。

假设一 假设日光温室中的初始害虫是在温室建成之前由外界迁入的或是土壤中存留的，温室建成后不再有害虫从外界迁入。

假设二 假设温室中的蔬菜在生长期内不会出现因数量不够而使害虫数下降的情况，因而对害虫的自然增长，我们可以参考使用 Malthus 模型。

假设三 假设以 1 天（24 小时）为单位时间。

假设四 假设每次使用农药的时间和使用农药后的药物有效时间之和 Δt 都较短，可以忽略不计。约定 $x(t)$ 表示 t 时刻（未用药时）的害虫密度；$X(t, u)$ 表示 t 时刻使用农药 u 千克，$t+\Delta t$ 时刻温室中害虫的残存密度。

假设五 假定每只害虫在单位时间内对蔬菜的破坏量是一个常数。

假设六 假设每次使用农药的设备使用费、人工费以及单位农药的费用都是固定的。

假设七 假定生长季节内任意时刻使用每千克杀虫药对虫害的杀死率为一个常数 R''。

2.变量及符号的说明

T：蔬菜的生长周期长度，单位是“天”。

$X(t)$：t 时刻没有使用农药时蔬菜上害虫的平均密度，单位是“只／公顷”。

$X(t, u)$：t 时刻使用农药 u 千克，在 $t+\Delta t$ 时刻蔬菜上残存害虫的平均密度，单位是“只／公顷”。

X_0：生长季节初期（t=0 时刻）蔬菜上害虫的平均密度，单位是“只／公顷”。

β：单位时间（24 小时）内害虫的自然增长率。

α：在生长期内任意时刻 1 公顷蔬菜使用 1 千克杀虫药对虫害的杀死率。

Q：每只害虫在单位时间内对蔬菜的破坏量，单位是“千克”。

D：蔬菜的单价，单位是“元”。

V：每次使用杀虫药设备的固定费用，单位是“元”。

W：杀虫药单价，单位是“元”。

R：每次使用杀虫药的人工费用，单位是“元”。

M: 蔬菜生长季节内单位面积每次农药用量不能超过的最大值，单位是“千克”。

4.4.4　模型建立

设在蔬菜生长季节[0，*T*]内杀虫的次数是 n 次，记每次杀虫的时刻依次为 $t_0,t_1,\cdots,t_{n-1}$，每平方米蔬菜用药剂量依次为 $u_0,u_1,\cdots,u_{n-1}$。我们的目标是确定 n 的值和用药时刻 $t_0,t_1,\cdots,t_{n-1}$ 以及单位面积用药剂量 $u_0,u_1,\cdots,u_{n-1}$，使得在一定条件下总费用最低。

1.害虫密度关于时间的变化规律

当在 t_{k-1} 时刻使用数量为 u_{k-1} 的杀虫药后，在时间区间 $(t_{k-1}+\Delta t,\ t_k)$（Δt 可以忽略不计）内不再使用杀虫药，由于温室建成后不再有害虫从外界迁入，所以蔬菜上害虫的平均密度 $X(t)$ 在 $t\in(t_{k-1},\ t_k)$ 时刻增长率 $\frac{\mathrm{d}X(t)}{\mathrm{d}t}$ 只与自然增长率 β 有关，于是得到微分方程

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t}=\beta X(t) \tag{4.16}$$

初始条件为在时刻 t_{k-1} 使用 u_{k-1} 千克的杀虫药后残存害虫的平均密度，即：

$$X(t+\Delta t)\Big|_{t=t_{k-1}}=X(t_{k-1},u_{k-1}),k=0,1,\cdots,n \tag{4.17}$$

解此初值问题得

$$X(t)=X(t_{k-1},u_{k-1})\mathrm{e}^{\beta(t-t_{k-1}-\Delta t)},t_{k-1}<t<t_k\ ,k=0,1,\cdots,n \tag{4.18}$$

在(4.17)和(4.18)式中，$t_{-1}=0,u_{-1}=0,t_n=T,X(t_{-1},u_{-1})=X(0)=X_0$，而 $X(t_{k-1},u_{k-1})$ 是在区间 (t_{k-1},t_k) 左端点时害虫的残存密度。

由假设 Δt 忽略不计，于是得到区间 (t_{k-1},t_k) 内害虫密度增长方程：

$$X(t)=X(t_{k-1},u_{k-1})\mathrm{e}^{\beta(t-t_{k-1})},t_{k-1}<t<t_k\ ,k=0,1,\cdots,n \tag{4.19}$$

2.杀虫药使用剂量对害虫密度的影响

根据假设，单位面积内使用每千克的杀虫药对害虫的杀死率为一个常数 α，在时刻 $t_k\in[0,T]$ 处使用数量为 $u(u\geq 0)$ 的杀虫药后残存害虫的平均密度应满足下面的偏微分方程初值问题：

$$\begin{cases}\dfrac{\partial X(t_k,u)}{\partial u}=-\alpha X(t_k)X(t_k,u)\\ X(t_k,0)=X(t_k)\end{cases}$$

解此初值问题得 $X(t_k,u)=X(t_k)e^{-\alpha u},\quad k=0,1,\cdots,n-1$　　（4.20）

在（4.19）式中令 $t\to t_k^-$ 得到

$$X(t_k)=X(t_{k-1},u_{k-1})\mathrm{e}^{\beta(t_k-t_{k-1})},k=0,1,\cdots,n \tag{4.21}$$

把（4.21）式代入（4.20）式得到用药后残存害虫的密度为：

$$X(t_k,u)=X(t_{k-1},u_{k-1})e^{\beta(t_k-t_{k-1})-\alpha u},\ k=0,1,\cdots,n-1 \tag{4.22}$$

在（4.22）式中取 $u=u_k$ 可得

$$X(t_k,u_k)=X(t_{k-1},u_{k-1})e^{\beta(t_k-t_{k-1})-\alpha u_k},\quad k=0,1,\cdots,n-1 \tag{4.23}$$

显然，残存害虫密度 $X(t_k,u_k)$ 和前次残存害虫密度 $X(t_{k-1},u_{k-1})$ 、两次用药时间间隔 t_k-t_{k-1} 以及本次用药量 u_k 都有关，而与其他因素无关。

3.蔬菜在生长期内总损失费最小的数学模型

在时间区间 $(t_{k-1},\ t_k)$ 内被害虫破坏的蔬菜损失费

$$E_k=DQS\int_{t_{k-1}}^{t_k}X(t)dt=DQS\int_{t_{k-1}}^{t_k}X(t_{k-1},u_{k-1})e^{\beta(t-t_{k-1})}dt$$

$$=\frac{DQS}{\beta}X(t_{k-1},u_{k-1})\left(e^{\beta(t_k-t_{k-1})}-1\right),\quad k=0,1,\cdots,n$$

于是在蔬菜生长期 $[0,T]$ 内被害虫破坏的蔬菜的总损失费

$$E_D=\sum_{k=0}^{n}E_k=\frac{DQS}{\beta}\sum_{k=0}^{n}X(t_{k-1},u_{k-1})\left(e^{\beta(t_k-t_{k-1})}-1\right) \tag{4.24}$$

农药使用的总费用为

$$E_Z=n(R+V)+SW\sum_{k=1}^{n}u_{k-1}$$

综上，我们得到蔬菜在生长期 $[0,T]$ 内总损失费的数学模型

$$\begin{cases}E(t_{k-1},u_{k-1},n)=E_Z+E_D=n[R+V]+SW\sum_{k=1}^{n}u_{k-1}+\dfrac{DQS}{\beta}\sum_{k=0}^{n}X(t_{k-1},u_{k-1})\left(e^{\beta(t_k-t_{k-1})}-1\right)\\ X(t_k,u_k)=X(t_{k-1},u_{k-1})e^{\beta(t_k-t_{k-1})-\alpha u_k},0\le u_k\le M,t_{k-1}<t_k,\quad k=0,1,\cdots,n-1\end{cases} \tag{4.25}$$

从而得到总损失费最小的数学模型

$$\begin{cases}E_{\min}(t_{k-1},u_{k-1},n)=\min\left\{n[R+V]+SW\sum_{k=1}^{n}u_{k-1}+\dfrac{DQS}{\beta}\sum_{k=0}^{n}X(t_{k-1},u_{k-1})\left(e^{\beta(t_k-t_{k-1})}-1\right)\right\},\\ X(t_k,u_k)=X(t_{k-1},u_{k-1})e^{\beta(t_k-t_{k-1})-\alpha u_k},0\le u_k\le M,t_{k-1}<t_k,k=0,1,\cdots,n-1\end{cases} \tag{4.26}$$

显然，此模型实际上是一个关于 t_{k-1} ，u_{k-1} ，n 的多元非线性规划。

4.4.5 模型简化

考虑到在生长期 $[0,T]$ 内的任意非用药时刻 t 处害虫数的自然增长率是常数 β ，我们可以采用从第一次用药时刻 t_0 开始周期性地使用杀虫药的策略。即每隔相同的时间使用相同剂量的农药，使得用药后害虫密度都下降为某一个限值 N 。这样的操作方式不仅可以简化模型（4.25），而且可以降低菜农使用农药的操作难度。因此值得我们进一步研究。

设使用农药的周期性时间间隔 $t_k-t_{k-1}=h,k=1,2,\cdots,n$ ，$t_n=T$ ， $T=nh+t_0$ 。

1.温室中害虫的初始密度大于某个限值 N

在蔬菜生长期的初始时刻，若满足 $X_0 = X(t_0) > N > 0$，则考虑先使用数量 $u_0 = \xi$ 的杀虫药（ξ 的具体数值待定），使得害虫密度降低到 $X(t_0, \xi) = N$。为简便起见，取第一次使用农药的时间为最初时刻，即令 $t_0 = 0$。经过一个周期也就是经过 h 天后，由（4.21）式得知，害虫密度在 t_1 时刻增长到 $X(t_1)=X(t_0, u_0)\mathrm{e}^{\beta(t_1-t_0)} = X(0,\xi)\mathrm{e}^{\beta h} = N\mathrm{e}^{\beta h}$，此时立即使用剂量为 u 的杀虫药，使害虫密度降低到 N，即

$$X(t_1, u) = X(t_0, \xi)\mathrm{e}^{\beta h-\alpha u} = N\mathrm{e}^{\beta h-\alpha u} = N$$

设以后单位面积内每次的用药量都为 u，即有 $u_1 = u_2 = \cdots = u_n = u$。经过 h 天后，在 t_2 时刻害虫密度增长到 $X(t_2) = X(t_1, u)\mathrm{e}^{\beta h} = N\mathrm{e}^{\beta h}$，这时使用剂量为 u 的杀虫药，使害虫密度降低到 N，即 $X(t_2, u) = X(t_1, u)\mathrm{e}^{\beta h-\alpha u} = N\mathrm{e}^{\beta h-\alpha u} = N$，根据公式（4.21）和公式（4.22）经过完全归纳得出

$$X(t_k) = X(t_{k-1}, u)\mathrm{e}^{\beta h} = N\mathrm{e}^{\beta h},\ X(t_k, u) = X(t_{k-1}, u)\mathrm{e}^{\beta h-\alpha u} = N\mathrm{e}^{\beta h-\alpha u} = N, k = 1, 2, \cdots, n$$

害虫密度将在每一周期末达到最大值，函数 $X(t)$ 从 t_0 开始呈现出周期性的变化。

在（4.24）中令 $u_{k-1} = u$，$X(t_{k-1}, u_{k-1}) = N$，$t_k - t_{k-1} = h$，$k = 1, 2, \cdots, n$，得到在蔬菜生长期内被害虫破坏的蔬菜的价值

$$E_D = \sum_{k=1}^{n} E_k = \frac{nDQSN}{\beta}\left(\mathrm{e}^{\beta h} - 1\right)$$

农药使用的总费用为

$$E_Z = (R + V)n + SW\left[(n-1)u + \xi\right]$$

因 $\beta h = \alpha u$，$n = \dfrac{\beta T}{\alpha u}$，$N = X_0\mathrm{e}^{-\alpha\xi}$，所以在周期性等量使用农药的情况下，模型（4.25）简化为

$$\begin{cases} E(u,\xi) = E_D + E_Z = \dfrac{DQSX_0T}{\alpha u}\mathrm{e}^{-\alpha\xi}\left(\mathrm{e}^{\alpha u} - 1\right) + (R+V)\dfrac{\beta T}{\alpha u} + \dfrac{SW}{\alpha}\left[\beta T - \alpha(u - \xi)\right] \\ h = \dfrac{\alpha u}{\beta},\ n = \dfrac{\beta T}{\alpha u},\ N = X_0\mathrm{e}^{-\alpha\xi} \end{cases} \tag{4.27}$$

总损失费最小的数学模型被简化为

$$E_{\min}(u,\xi) = \min_{0<\alpha<1} E(u,\xi),\ h = \frac{\alpha u}{\beta},\ n = \frac{\beta T}{\alpha u},\ N = X_0\mathrm{e}^{-\alpha\xi} \tag{4.28}$$

2.温室中害虫的初始密度不大于某个限值 N

若在蔬菜生长期的初始时刻 $X_0 = X(t_0) \le N$，我们可暂不杀虫（虫害不严重，没有必要立即杀），采用经过一段时间 h_0，待害虫密度增长到一定水平 N 后再采用上

述周期性杀灭的办法，h_0 的值可以通过关系式 $X(h_0)=X_0\mathrm{e}^{\beta h_0}=N$ 求得，此时函数 $X(t)$ 也呈现周期性的变化。

在 $X_0\le N$ 的情况下，总时间 $T=(n+1)h+h_0$，仍将单位面积内每次农药的使用量记为 u，与前面 $X_0>N$ 时的讨论相同。

因 $\beta h=\alpha u$，$n+1=\dfrac{\beta(T-h_0)}{\alpha u}$，所以在时间区间 $[t_{k-1},\ t_k]$ 内被害虫破坏的蔬菜价值

$$E_k=\frac{DQS}{\beta}X(t_{k-1},u)(\mathrm{e}^{\beta h}-1)=\frac{DQSN}{\beta}(\mathrm{e}^{\beta h}-1),\quad k=1,2,\cdots,n$$

因而在蔬菜生长期 $[0,T]$ 内被害虫破坏的蔬菜总价值

$$E_D=DQS\int_0^{t_0}X(t)\mathrm{d}t+\sum_{k=1}^{n}E_k=DQS\left(\int_0^{h_0}X_0\mathrm{e}^{\beta t}\mathrm{d}t+\int_{h_0}^{t_0}X_0\mathrm{e}^{\beta t}\mathrm{d}t\right)+\frac{nDQSN}{\beta}(\mathrm{e}^{\beta h}-1)$$
$$=\frac{DQSX_0}{\beta}(\mathrm{e}^{\beta h_0}-1)+\frac{DQSN}{\beta}(n+1)(\mathrm{e}^{\beta h}-1)$$

农药使用总费用 $E_Z=n(R+V+SWu)$。于是蔬菜在生长期内总损失费

$$\begin{aligned}E(u,h_0)&=E_Z+E_D\\&=n(R+V+SWu)+\frac{DQSX_0}{\beta}(\mathrm{e}^{\beta h_0}-1)+\frac{DQSN}{\beta}(n+1)(\mathrm{e}^{\beta h}-1)\\&=\frac{\beta(T-h_0)}{\alpha u}\left[R+V+SWu+\frac{DQSX_0}{\beta}\mathrm{e}^{\beta h_0}(\mathrm{e}^{\alpha u}-1)\right]\\&\quad+\frac{DQSX_0}{\beta}(\mathrm{e}^{\beta h_0}-1)-(R+V+SWu)\end{aligned}\tag{4.29}$$

总损失费最小的数学模型

$$E_{\min}(u,h_0)=\min_{0\le u\le M,0<h_0<T}E(u,h_0),\ h=\frac{\alpha u}{\beta},\ n=\frac{\beta(T-h_0)}{\alpha u}-1,\ N=X_0\mathrm{e}^{\beta h_0}\tag{4.30}$$

4.4.6 模型求解

模型的求解不难实现，只要给出各个参数的值，应用数学软件即可求出模型（4.28）和（4.30）的最优数值解，哪个模型的目标函数值较小就选用哪个操作方案。如果模型（4.28）比模型（4.30）的值小，则用药策略是：在最初时刻先用药 ξ 千克，然后每隔时间 h 天用剂量为 u 千克的药杀一次虫；如果模型（4.28）比模型（4.30）的值大，则用药策略是：在时刻 h_0+h 天第一次用药 u 千克，然后每隔固定周期时间 h 天用固定剂量 u 千克的药杀虫。

例如，在本节开始的问题中假设：

S=1, T=100, M=1, X_0=7500, a=0.5, Q=0.001, β=0.02, V=10, D=10, W=12, R=15。

将这些数据代入（4.27）得：

$$E(u,\xi)=\frac{15000}{u}e^{-0.5\xi}\left(e^{0.5u}-1\right)+\frac{100}{u}-12u+12\xi+48,(u,\xi)\in(0,1]\times(0,1] \quad (4.31)$$

这个二元函数在区域$(0,1]\times(0,1]$上连续，因而在此区域的任意子闭区域上必能取到最小值，运用 Mathcad 的求导法及 Given-Find 结构找不到稳定点，说明极小值点在区域边界上。不妨取$u=1$，则有

$$f(\xi)=E(1,\xi)=15000\left(e^{0.5}-1\right)e^{-0.5\xi}+12\xi+136,\quad \xi\in(0,1] \quad (4.32)$$

因$\frac{d}{d\xi}f(\xi)=-7500\left(e^{0.5}-1\right)e^{-0.5\xi}+12<0\,,\xi\in(0,1]$，所以$f(\xi)$在$(0,1]$上严格递减（如图 4-3）。

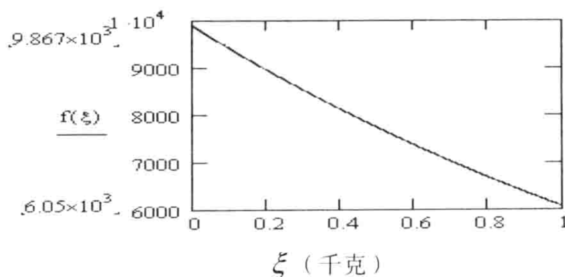


图 4-3 模型（4.32）的单调性直观图

因此有最小值$f(1)=E(1,1)=6050$。从而求得$E(u,\xi)$在区域$(0,1]\times(0,1]$上的最小值为$E(1,1)=6050$（参看图 4-4）。

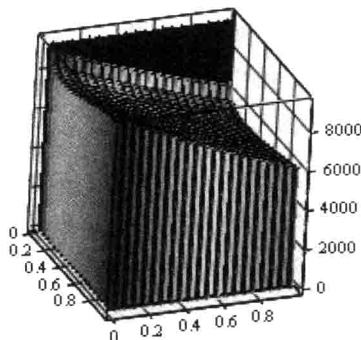


图 4-4 模型（4.31）的最优数值解直观图

于是得到模型（4.28）的最优数值解：$\xi=1$千克，$u=1$千克，最小损失费$E=6050$元。同时还可求得农药使用方案的其他参数值 n=4 次，h=25 天。

同样，将这些数据代入（4.29）得

$$E(u,h_0)=\frac{100-h_0}{25u}\left[25+12u+3750e^{0.02h_0}\left(e^{0.5u}-1\right)\right]+3750\left(e^{0.02h_0}-1\right)$$
$$-(25+12u),\ (u,h_0)\in(0,1]\times[0,100) \quad (4.33)$$

运用 MATHCAD 软件程序求得此模型下的最小损失费是 $E(0.23,0)$=8403 元.（参看图 4-5）。

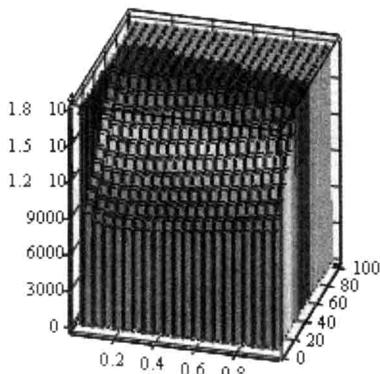


图 4-5 模型（4.33）的最优数值解直观图

由于 $E(1,1)=6050<8403=E(0.23,0)$，所以我们选择方案（4.28），最小损失费用应为模型（4.28）的解 6050 元。由此我们得出该日光温室农药的使用策略如下：在最初时刻使用农药剂量 1 千克，此后每隔 25 天再使用一次杀虫药，每次的用药剂量为 1 千克，在整个蔬菜生长期内总损失费用为 6050 元。

模型（4.28）和模型（4.30）代表了两种不同的用药策略，仔细分析可知，决定使用哪个策略的关键因素是初始害虫的密度。比如在 $X_0=9000$ 只时使用模型（4.28）的策略比模型（4.30）要更加节省费用，但在 $X_0=750$ 只时使用模型（4.28）的策略比模型（4.30）要多一些费用。可以用数值方法求得导致使用策略在两个模型之间变换的 X_0 的临界值。

4.4.7 模型评价

为了简化模型，我们把不定期地多次使用杀虫药的多元非线性规划问题转化成了周期性使用杀虫药的两个二元非线性规划问题，这使模型求解可以直接采用数学软件 Mathcad 来实现。比较两个简化模型的值，即可导出最后的结论。

作为对比，考察几种较简单的极端的情况。

如果不使用杀虫药，则被害虫损坏的蔬菜总损失费用

$$E=DQSX_0\int_0^{100}e^{\beta t}\,dt=\frac{DQSX_0}{\beta}\left(e^{100\beta}-1\right)$$

$$=\frac{10\times0.001\times1\times7500}{0.02}\left(e^{100\times0.02}-1\right)=23960\text{ 元}$$

这一损失费用远远高于 6050 元。

如果只使用一次杀虫药，经与第 4 部分类似的分析易得到如下的模型

$$E(t,u)=\frac{DQSX_0}{\beta}\left[\mathrm{e}^{\beta t-\alpha u}\left(\mathrm{e}^{\beta(T-t)}-1\right)+\mathrm{e}^{\beta t}-1\right]$$
$$+R+V+SWu\,,0<u\le M\,,\ 0<t<T \tag{4.34}$$

多次用药模型是一种容易用计算机程序实现的优化模型，该模型不仅在蔬菜生产中具有很强的应用价值，而且可推广到传染病的控制、保护生态平衡等其他决策问题中去，但该模型也有一些缺陷，例如，“在单位面积内使用每单位的农药，其杀死害虫的比例为一个常数α”的假设仅在害虫密度比较高的时候才近似成立，将模型结果用于密度较低的情形会产生较大的模型误差。在这种情形下应当如何操作，还有待于进一步的分析和研究。

习题四

1.设z是和平国家，x与y是好战国，则三国的军备x,y,z满足

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=-ax+ky+kz+g_1\\[2mm]\dfrac{\mathrm{d}y}{\mathrm{d}t}=kx-ay+kz+g_2\\[2mm]\dfrac{\mathrm{d}z}{\mathrm{d}t}=-az+g_3\end{cases}$$

试讨论解$x(t),y(t),z(t)$的稳定性。

2. A方反坦克导弹与B方坦克之间进行战斗。它们之间的起始距离是d，坦克以等速向反坦克导弹驶进。设反坦克导弹一次发射击毁坦克的概率为$p_A(t)=0.6+0.05t$，坦克一次发射击毁反坦克导弹的概率为$p_B(t)=0.4+0.3t$。试确定双方的最优开火距离。

3.方程组

$$\begin{cases}\dot{x}=-ay\\ \dot{y}=-by-cxy\end{cases}$$

是正规军对游击队作战的一个模型，其中$y(t)$是游击队，试求其轨线，分析胜负。

4.方程组

$$\begin{cases}\dot{x}=-ay-cxy\\ \dot{y}=-by-dxy\end{cases}$$

是游击队对游击队作战的数学模型，试求其轨线，分析胜负。

5.地球人口极限为 100 亿，1974 年世界人口为 40 亿，试用阻滞增长（Logistic）模型

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t}=0.029N(t)-2.941\times10^{-12}N(t)^2$$

求世界人口达到100亿需要多少年。

6.令Y为全国总产值，I为投资，G为政府开支，C是消费，对物资和服务的需求D为$D=C+I+G$，设$C=(1-s)Y$，s是储蓄系数，且设产量对过度需求的响应是

$$\frac{\mathrm{d}Y}{\mathrm{d}t}=l(D-Y),$$

其中，l是常数。投资的响应按下式计算：

$$\frac{\mathrm{d}I}{\mathrm{d}t}=m\left(\alpha\frac{\mathrm{d}Y}{\mathrm{d}t}-I\right),$$

其中，α与m为常数，G也为常数，试建立$Y(t)$满足的微分方程，且用它分析国民生产总值的变化情况。

第 5 章　插值与拟合建模

在实际中，常常要处理由实验或测量所得到的一批离散数据。插值与拟合方法就是要通过这些数据去确定某一类已知函数的参数或寻找某个近似函数，使所得到的近似函数与已知数据有较高的拟合精度。若要求这个近似函数（曲线或曲面）经过已知的所有数据点，则称此类问题为插值问题。

当所给的数据较多时，用插值方法得到的插值函数会很复杂，所以通常插值方法用于数据较少的情况。但通常情况下，数据一般都是由观测或实验得到的，往往会带有一定的随机误差，因而要求近似函数（曲线或曲面）通过所有的数据点也是不必要的。

若不要求近似函数通过所有数据点，而是要求它能较好地反映数据的整体变化趋势，则解决这类问题的方法称为数据拟合。

插值与拟合都是根据实际中一组已知数据来构造一个能够反映数据变化规律的近似函数的方法，由于对近似要求的准则不同，因此二者在数学方法上有很大的差异。

插值与拟合的方法很多，本章主要介绍一般插值方法、样条插值方法和最小二乘拟合方法。

5.1　插值方法

已知某函数 $y=f(x)$（未知）的一组观测（或实验）数据 $(x_i, y_i)\ (i=1,2,\cdots,n)$，要寻求一个函数 $\phi(x)$，使 $\phi(x_i)=y_i\ (i=1,2,\cdots,n)$，则 $\phi(x)\approx f(x)$。

具体地讲，实际中常常在不知道函数 $y=f(x)$ 的具体表达式的情况下，对于 $x=x_i$ 有实验测量值 $y=y_i\ (i=0,1,2,\cdots,n)$，寻求另一函数 $\phi(x)$ 使满足：

$$\phi(x_i)=y_i=f(x_i)\ (i=0,1,2,\cdots,n)$$

称此问题为插值问题，并称函数 $\phi(x)$ 为 $f(x)$ 的插值函数，$x_0, x_1, x_2, \cdots, x_n$ 称为插值节点，$\phi(x_i)=y_i\ (i=0,1,2,\cdots,n)$ 称为插值条件，即 $\phi(x_i)=y_i=f(x_i)\ (i=0,1,2,\cdots,n)$，

则 $\phi(x) \approx f(x)$。

5.1.1 线性插值

设有 $x_0 = a < x_1 < x_2 < \cdots < x_n = b$,(x_i, y_i) 是已知的 $(n+1)$ 个节点，求任一插值点 $x^*(\neq x_i)$ 的插值函数值 y^*。

一般地，插值函数要求尽可能地简单、实用。其中最简单的是线性插值——用这些节点的连接折线表示，记为 $I_n(x)$。

$$I_n(x) = \sum_{j=0}^{n} y_j l_j(x)$$

$$l_j(x) = \begin{cases} \dfrac{x - x_{j-1}}{x_j - x_{j-1}} & (x_{j-1} \le x \le x_j)(j \neq 0) \\[2ex] \dfrac{x - x_{j+1}}{x_j - x_{j+1}} & (x_j \le x \le x_{j+1})(j \neq n) \\[2ex] 0 & (\text{其他}) \end{cases}$$

在用函数表示插值计算时，分段线性插值即可。

5.1.2 Lagrange 插值

设函数 $y = f(x)$ 在 $n+1$ 个相异点 $x_0, x_1, x_2, \cdots, x_n$ 上的函数值为 $y_0, y_1, y_2, \cdots, y_n$，要求一个次数不超过 n 次的代数多项式

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

使在节点 x_i 上有 $p_n(x_i) = y_i \ (i = 0, 1, 2, \cdots, n)$ 成立，称之为 n 次代数插值问题。$p_n(x)$ 称为插值多项式。可以证明 n 次代数插值问题的解是唯一的。

事实上，可以得到

$$p_n(x) = \sum_{j=0}^{n} \left[\prod_{\substack{i=0 \\ i \neq j}}^{n} \left(\frac{x - x_i}{x_j - x_i} \right) \right] y_j$$

当 $n=1$ 时，有两点一次线性插值多项式：

$$p_2(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1$$

当 $n=2$ 时，有三点二次抛物线插值多项式：

$$p_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2$$

5.1.3 Newton 插值

1. Newton 插值的基本思想

由于 $y = f(x)$ 关于两节点 x_0, x_1 的线性插值多项式为

$$p_1(x)=f(x_0)+\frac{f(x_1)-f(x_0)}{x_1-x_0}(x-x_0)=p_0(x_0)+\frac{f(x_1)-f(x_0)}{x_1-x_0}(x-x_0)$$

假设满足插值条件 $f(x_i)=y_i=p_2(x_i)\,(i=0,1,2)$ 的二次插值多项式一般形式为

$$p_2(x)=c_0+c_1(x-x_0)+c_2(x-x_0)(x-x_1)$$

由插值条件可得

$$\begin{cases} c_0=f(x_0) \\ c_0+c_1(x_1-x_0)=f(x_1) \\ c_0+c_1(x_2-x_0)+c_2(x_2-x_0)(x_2-x_1)=f(x_2) \end{cases}$$

可以解出

$$\begin{cases} c_0=f(x_0) \\ c_1=\dfrac{f(x_1)-f(x_0)}{x_1-x_0} \\ c_2=\dfrac{\dfrac{f(x_2)-f(x_1)}{x_2-x_1}-\dfrac{f(x_1)-f(x_0)}{x_1-x_0}}{x_2-x_0} \end{cases}$$

所以

$$p_2(x)=c_0+c_1(x-x_0)+c_2(x-x_0)(x-x_1)=p_1(x)+c_2(x-x_0)(x-x_1)$$

用类似的方法，可以得到三次插值多项式等，按这种思想可以得到一般的 Newton 插值公式。

2.函数的差商及其性质

对于给定的函数 $f(x)$，用 $f(x_0,x_1,\cdots,x_n)$ 表示关于节点 x_0，x_1，x_2，$\cdots,x_n$ 的 n 阶差商，则有

一阶差商　$f(x_0,x_1)=\dfrac{f(x_1)-f(x_0)}{x_1-x_0}$，$f(x_1,x_2)=\dfrac{f(x_2)-f(x_1)}{x_2-x_1}$

二阶差商　$f(x_0,x_1,x_2)=\dfrac{f(x_1,x_2)-f(x_0,x_1)}{x_2-x_0}$

……

n 阶差商　$f(x_0,x_1,\cdots,x_n)=\dfrac{f(x_1,x_2,\cdots,x_n)-f(x_0,x_1,\cdots,x_{n-1})}{x_n-x_0}$

差商有下列性质：

（1）差商的可加性 $f(x_0, x_1, \cdots, x_n) = \sum_{k=0}^{n} \frac{f(x_k)}{\prod_{\substack{j=0 \\ j \neq k}}^{n} (x_k - x_j)}$；

（2）差商的对称性 在 $f(x_0, x_1, \cdots, x_n)$ 中任意调换 x_i, x_j 的次序其值不变。

3.Newton 插值公式

由插值条件可以相应地确定函数差商，于是有一般的 Newton 插值公式

$$\begin{aligned} p_n(x) &= f(x_0) + f(x_0, x_1)(x - x_0) + f(x_0, x_1, x_2)(x - x_0)(x - x_1) + \cdots \\ &\quad + f(x_0, x_1, \cdots, x_n)(x - x_0)(x - x_1)\cdots(x - x_{n-1}) \\ &= p_{n-1}(x) + f(x_0, x_1, \cdots, x_n)(x - x_0)(x - x_1)\cdots(x - x_{n-1}) \end{aligned}$$

可以证明：其余项为

$$R(x) = f(x_0, x_1, \cdots, x_n, x)(x - x_0)(x - x_1)\cdots(x - x_{n-1})(x - x_n)$$

特别地，一次插值公式为

$$p_1(x) = f(x_0) + f(x_0, x_1)(x - x_0)$$

二次插值公式为

$$\begin{aligned} p_2(x) &= f(x_0) + f(x_0, x_1)(x - x_0) + f(x_0, x_1, x_2)(x - x_0)(x - x_1) \\ &= p_1(x) + f(x_0, x_1, x_2)(x - x_0)(x - x_1) \end{aligned}$$

实际上，Newton 插值公式是 Lagrange 插值公式的一种变形，二者是等价的。

5.1.4 样条插值

样条曲线来自工程，普遍使用的是三次样条插值，原因是三次样条插值符合力学要求，且有较好的分析性质。

三次样条插值函数 $S(x)$，$a \le x \le b$，要求：

（1）在 $[x_{i-1}, x_i]$ 上为三次多项式；

（2）在 $[a, b]$ 上二阶导数连续；

（3）$S(x_i) = y_i, i = 0, 1, \cdots, n$。

注意：

（1）在数据被认为是严格精确时，用样条插值；

（2）用 Lagrange 插值，次数不宜太高（即数据点过多），应在 8 次以下；

（3）若数据不太准，或是统计结果，用回归方法。

5.2　拟合方法

5.2.1　线性最小二乘法

设有一组数据$(x_i, y_i), i=1,2,\cdots,n$，求一直线与$(x_i, y_i)$尽可能接近，即求一条一元回归直线$y=a+bx+\varepsilon$，使$y^*=a+bx$与$(x_i, y_i)$尽可能接近。

一般地，求a,b使$Q_{\min}=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$，即求$Q(a,b)=\sum_{i=1}^{n}(y_i-a-bx_i)^2$的最小值点。

设满足上述函数最小值的点为$(\hat{a},\hat{b})$，则称$\hat{y}=\hat{a}+\hat{b}x$为一元线性回归方程（或回归直线），$\hat{a},\hat{b}$称为回归系数。这种求回归方程的方法称为最小二乘法。

所求可由数学分析或高等数学中求极值的方法得到。

由$\dfrac{\partial Q}{\partial a}=0=\dfrac{\partial Q}{\partial b}$解得，

$$\hat{b}=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2};\quad \hat{a}=\overline{y}-\hat{b}\overline{x}$$

其中

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

记　$L_{xx}=\sum_{i=1}^{n}(x_i-\overline{x})^2, L_{xy}=\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})$，则

$$\hat{b}=\frac{L_{xy}}{L_{xx}},\ \hat{a}=\overline{y}-\hat{b}\overline{x}$$

用数学软件可以方便地求出$\hat{a},\hat{b}$，基本命令为$Fit\left[d,\{1,x\},x\right]$。

$\hat{a}$和$\hat{b}$的置信区间：设$s=\dfrac{SSE}{n-2}$，$\hat{\sigma}_{\hat{a}}=\sqrt{\dfrac{1}{n}+\dfrac{\overline{x}^2}{L_{xx}}}s$，$\hat{\sigma}_{\hat{b}}=\dfrac{s}{\sqrt{L_{xx}}}$，则

$\hat{a}$的置信区间为

$$\left(\hat{a}-\sqrt{F_{1-\alpha}(1,n-2)}\hat{\sigma}_{\hat{a}},\ \hat{a}+\sqrt{F_{1-\alpha}(1,n-2)}\hat{\sigma}_{\hat{a}}\right)$$

$\hat{b}$的置信区间为

$$\left(\hat{b}-\sqrt{F_{1-\alpha}(1,n-2)}\hat{\sigma}_{\hat{b}},\ \hat{b}+\sqrt{F_{1-\alpha}(1,n-2)}\hat{\sigma}_{\hat{b}}\right)$$

如何来检验所得回归方程是否符合实际呢?

5.2.2　相关性和显著性检验

设有一组数据$(x_i, y_i), i=1,2,\cdots,n$，若这组数据有较好的线性关系，则用直线$y=a+bx$拟合；若不具有良好的线性关系，则用线性回归效果不理想。当用这一组

数据做线性回归时，先要做统计检验分析。统计检验常用的方法有相关系数检验和 F 检验。

（1）相关系数检验

$$r=\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}$$

r 称为 x 与 y 的相关系数，它反映了 x 与 y 的线性关系和程度。可以证明 $|r|\le 1$。

$r=\pm 1$ 表示有精确的线性关系。如 $y_i=a+bx_i$，则 $b>0$ 时 $r>0$，表示正线性相关；$b<0$ 时 $r<0$，表示负线性相关。下面就 $b>0$ 的情形给出证明， $b<0$ 情形的证明同理。

$$\begin{aligned}r&=\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}\\&=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(a+bx_i-a-b\bar{x})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(a+bx_i-a-b\bar{x})^2}}\\&=1\end{aligned}$$

如图 5-1 所示，一般地，$|r|$ 越小，相关性越差。

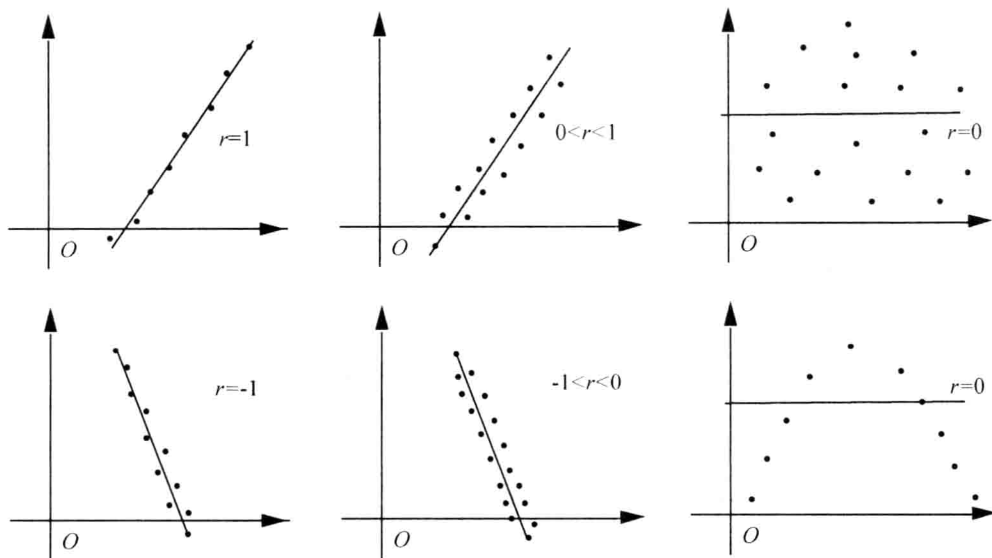


图 5-1 相关系数的几何意义

由此，可根据 r 的大小查相关系数表(见文献[2]的附录)判断数据的相关性：$|r|$ 小于 5%对应的值时，称为无线性关系；$|r|$ 大于 5%而小于 1%对应的值时，称为有显

著的线性关系；$|r|$大于 1%对应的值时，称为有极显著的线性关系。

例如，$n=6$, 5% 对应于 0.811，1%对应于 0.917，于是，$|r|$在 0.811 与 0.917 之间时，认为有显著的线性关系；$|r|$在 0.917 以上时，认为有极显著的线性关系；$|r|$在 0.811 以下时，认为无线性关系。

此外，还可使用 F 检验。

（2）F 检验

用 F 检验检查 n 个观察值之间的差异，可用

$$S_{总}=\sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2=L_{yy}$$

衡量，称为总离差平方和。

但是，$S_{总}$ 的大小与 n 及 y_i 偏离 $\overline{y}$ 的大小均有关，因此无法衡量。通常要设法化为某个统计量有关的数据，从而，可以用某个概率分布来衡量。

作如下处理：

由　$y_i-\overline{y}=y_i-\hat{y}_i+\hat{y}_i-\overline{y}$ 得

$$S_{总}=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2+\sum_{i=1}^{n}\left(\hat{y}_i-\overline{y}\right)^2+2\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)\left(\hat{y}_i-\overline{y}\right)$$

可以证明

$$\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)\left(\hat{y}_i-\overline{y}\right)=0$$

$$S_{总}=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2+\sum_{i=1}^{n}\left(\hat{y}_i-\overline{y}\right)^2=SSE+SSR$$

其中

$$SSE=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2\text{；}SSR=\sum_{i=1}^{n}\left(\hat{y}_i-\overline{y}\right)^2$$

称为残差平方和。

$$SSE=bL_{xx}\text{；}SSR=L_{yy}-bL_{xx}=L_{yy}-SSR$$

利用方差理论可知，SSR 的自由度为 1（一个变量），SSE 的自由度为 n-2。

定义

$$F=\frac{SSR/1}{SSE\sqrt{n-2}}$$

若 x 与 y 有线性关系，则 F 服从第一自由度 1、第二自由度为 n-2 的 F 分布，记为 $F(1,n-2)$。

按概率有关理论，可确定相应临界值 c，即　$P(F>c)=\alpha$，从而得到 F 检验，可由表 5-1 进行。

表 5-1

方差来源	平方和	自由度	$F_{比}$	显著性
回归	$SSR=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$	1	$F=\dfrac{SSR/1}{SSE/(n-2)}$	
残差	$SSE=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$	n-1-1		
总方差	$S_\gamma=\sum_{i=1}^{n}(y_i-\bar{y})^2$	n-2+1=n-1	$F_{1-\alpha}(1,n-2)$	

（1）$F_{0.95}(1,n-2)<F<F_{0.99}(1,n-2)$，线性关系显著；

（2）$F_{0.99}(1,n-2)<F$，线性关系极其显著；

（3）$F<F_{0.95}(1,n-2)$，无线性关系。

预测的精度δ：对于$0<\alpha<1$，有

$$\delta=\sqrt{F_{1-\alpha}(1,n-2)\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}}\right]}\cdot\frac{SSE}{n-2}$$

则对$x=x_0,\hat{y}_0=\hat{y}(x_0)$的预测区间为$(\hat{y}_0-\delta,\hat{y}_0+\delta)$。

当n较大时，有

$$1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}}\approx 1,\quad \delta\approx F_{1-\alpha}(1,n-2)$$

5.2.3 可化为一元线性回归的非线性回归

一般地，非线性回归无一般方法，但有些非线性回归问题可化为线性回归问题。这里介绍两类：

（1）$y=\beta_0+\beta_1\mathrm{e}^x$，$y=\beta_0+\beta_1x^2$，$y=\beta_0+\ln x$。

即自变量的指数函数、幂函数、对数函数与y有线性关系。

（2）$y=\beta_0\mathrm{e}^{\beta_1x}$，$y=\dfrac{1}{\beta_0+\beta_1\mathrm{e}^x}$，$y=\beta_0x^{\beta_1}$，$y=\dfrac{x}{\alpha x+\beta}$。

对于$y=\beta_0+\beta_1\mathrm{e}^x$，令$x'=\mathrm{e}^x,y'=y$，则$y'=\beta_0+\beta_1x'$化为一元线性回归。

对于$y=\beta_0\mathrm{e}^{\beta_1x}$，$\ln y=\ln\beta_0+\beta_1x$，令$x'=x,y'=\ln y,y'=\ln\beta_0+\beta_1x'$化为一元线性回归。其余可类推。

5.3 山东省职工平均工资预测模型

5.3.1 问题的提出

预测是指运用现有的方法，从已知事件推知未知事件。预测理论作为通用的方

法论，既可以应用于研究自然现象，也可以应用于研究社会现象。预测的本质是找到事物发展的趋势。所谓趋势就是众多“规律”中的一种表现形式。它描述的是一种渐进、连续发生的规律。趋势是事物明确的、可预见的发展方向。研究趋势是去观测一系列连续发生的事件，从中归纳出发展方向。而预测趋势就是去分析未来某个时间段内某个事物将会产生什么样的方向性变化。

试根据表 5-2 中的数据建立预测山东省职工年平均工资的数学模型，并由此预测 2011 年至 2035 年山东省职工的年平均工资。

表 5-2　山东省职工年平均工资统计表(单位：元)

年份	平均工资	年份	平均工资	年份	平均工资
1978	566	1989	1920	2000	8772
1979	632	1990	2150	2001	10007
1980	745	1991	2292	2002	11374
1981	755	1992	2601	2003	12567
1982	769	1993	3149	2004	14332
1983	789	1994	4338	2005	16614
1984	985	1995	5145	2006	19228
1985	1110	1996	5809	2007	22844
1986	1313	1997	6241	2008	26404
1987	1428	1998	6854	2009	29688
1988	1782	1999	7656	2010	32074

数据来源：山东省统计年鉴 2011.

5.3.2　模型建立

假设未来 30 年中国经济稳步发展，工资增长是随着时间连续变化的，即工资是时间的连续函数。

根据表 5-2 中的数据构造向量；

$$X=\begin{pmatrix}t_1 & t_2 & \cdots & t_{33}\end{pmatrix}=\begin{pmatrix}1978 & 1979 & \cdots & 2010\end{pmatrix};$$

$$Y=\begin{pmatrix}w_1 & w_2 & \cdots & w_{33}\end{pmatrix}=\begin{pmatrix}566 & 632 & \cdots & 32074\end{pmatrix};$$

利用 Mathcad 软件作出点 (t_i, w_i) $(i=1,2,\cdots,33)$，得山东省职工年平均工资散点图（见图 5-2）。

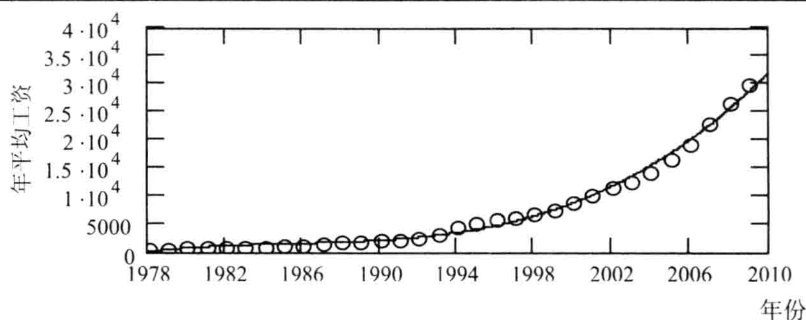


图 5-2 山东省职工年平均工资散点图与曲线拟合图

用 Mathcad 软件对这 33 个散点进行多项式拟合实验，发现用回归式抛物线拟合效果最佳，因此采用回归式抛物线进行拟合。

设拟合多项式为 $f(x)=ax^3+bx^2+cx+d$，其中 $a>0, a,b,c,d$ 为参数，我们选取 4 个点（1980, 745），（1990, 2150），（2000, 8772），（2010, 32074）以确定参数 a, b, c, d 的值。

利用 Mathcad 的 Given-Find 结构求解三元一次方程组，程序及结果如下：

$$
\begin{aligned}
&Given\\
&1980^3\cdot a+1980^2\cdot b+1980\cdot c+d=745\\
&1990^3\cdot a+1990^2\cdot b+1990\cdot c+d=2150\\
&2000^3\cdot a+2000^2\cdot b+2000\cdot c+d=8772\\
&2010^3\cdot a+2010^2\cdot b+2010\cdot c+d=32074
\end{aligned}
$$

$$
Find \to \begin{pmatrix} \dfrac{3821}{2000} \\ \dfrac{-56898}{5} \\ \dfrac{451874103}{20} \\ -14953001528 \end{pmatrix}
$$

所以

$$
f(x)=\frac{3821}{2000}x^3-\frac{56898}{5}x^2+\frac{451874103}{20}x-14953001528 \qquad (5.1)
$$

拟合效果参见图 5-2。

（5.1）式两端对 x 求导，再除以 $f(x)$ 得 $\dfrac{\mathrm{d}f(x)}{\mathrm{d}x}=g(x)f(x)$，其中 $g(x)$ 是 $f(x)$ 的变化率，且

$$
g(x)=\frac{3\left(3.821x^2-15172.8x+15062470.1\right)}{3.821x^3-22759.2x^2+45187410.3x-29906003056}
$$

显然 $\lim_{x\to\infty} g(x)=0$。这表明山东省职工年平均工资增长率 $g(x)$ 从某一时刻开始随着时间的推移逐渐递减并最终稳定于 0(见图 5-3)。这是符合实际的，因为工资不能无限增长，因此模型(5.1)是一种阻滞增长模型。另一方面从图 5-2 看拟合效果非常好，所以能够用模型(5.1)预测未来山东省职工的年平均工资。

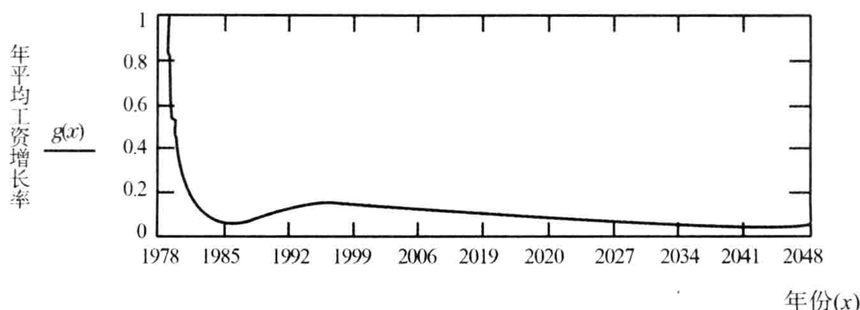


图 5-3　山东省职工年平均工资增长率曲线

5.3.3　山东省职工年平均工资预测

我们用模型(5.1)预测山东省 2011 年至 2035 年职工的年平均工资，结果如下：

$f(2011)=3.576\times10^4$　　$f(2024)=1.145\times10^5$

$f(2012)=3.974\times10^4$　　$f(2025)=1.234\times10^5$

$f(2013)=4.403\times10^4$　　$f(2026)=1.327\times10^5$

$f(2014)=4.863\times10^4$　　$f(2027)=1.424\times10^5$

$f(2015)=5.356\times10^4$　　$f(2028)=1.526\times10^5$

$f(2016)=5.883\times10^4$　　$f(2029)=1.634\times10^5$

$f(2017)=6.445\times10^4$　　$f(2030)=1.746\times10^5$

$f(2018)=7.043\times10^4$　　$f(2031)=1.863\times10^5$

$f(2019)=7.678\times10^4$　　$f(2032)=1.985\times10^5$

$f(2020)=8.352\times10^4$　　$f(2033)=2.113\times10^5$

$f(2021)=9.065\times10^4$　　$f(2034)=2.246\times10^5$

$f(2022)=9.819\times10^4$　　$f(2035)=2.385\times10^5$

$f(2023)=1.062\times10^5$

5.4　中国人口增长率模型

5.4.1　问题的提出

中国是人口大国，如何控制人口发展速度，关系到中华民族的存亡。因此，进

行人口预测是有效地控制人口发展与资源关系不可缺少的手段之一，同时也是合理地调整人口控制方案的决策依据。试从我国六次人口普查数据和人口增长的特点出发，运用数学建模的思想方法建立我国人口净增长率的数学模型，并由此对我国人口发展的中短期趋势做出预测。

5.4.2 模型的条件假设

要构建出较理想的数学模型，必须提出下列合理的假设；

假设一　不考虑意外灾难等因素对人口变化的影响；

假设二　不考虑我国人口同境外人口之间的迁入与迁出问题；

假设三　假设短期内我国的计划生育政策不变；

假设四　假设上一次人口普查到下一次人口普查这一段时间内，人口年净增长率取这一段时间内的平均值。

5.4.3 模型建立

我们将时间作为控制变量，人口净增长率作为状态变量，确定它们之间的函数关系。设 t 介于两个相邻的普查年份之间，根据中华人民共和国国家统计局人口普查公报记载的全国历次人口普查数据，计算人口年平均净增长率 $P(t)$，见表 5-3。

表 5-3　中国 6 次人口普查数据分析（人口总数单位：万人）

序号	普查年份	人口总量	增长量	净增长率	年均净增长率
1	1953	60193.8	—	—	—
2	1964	72307	12113.2	0.201	0.017
3	1982	103188.3	30881.3	0.427	0.02
4	1990	116001.7	12813.4	0.124	0.015
5	2000	129533	13531.3	0.117	0.011
6	2010	137053.7	7520.7	0.0581	0.0057

由于我国前三次人口普查并非每十年一次，时间间隔不相等，这给研究问题带来了不便。但假设了两个相邻普查年份之间各年的人口增长率取这一年段的平均值，所以选取下列五个点：

（1963, 0.017），（1974, 0.02），（1985, 0.015），（1996, 0.011），（2007, 0.0057）

构造一元线性回归

$$X=(1963 \quad 1974 \quad 1985 \quad 1996 \quad 2007)$$

$$Y=(0.017 \quad 0.02 \quad 0.015 \quad 0.011 \quad 0.0057)$$

利用 Mathcad 软件计算相关系数得

$$corr(X,Y)=-0.899$$

由于$0.878<\left|corr(X,Y)\right|<0.959$，所以 X，Y 之间存在显著的线性关系，也就是年份 t 和人口净增长率 $P(t)$ 具有良好的线性关系。不妨设 $P(t)\approx at+b$，其中 a,b 是参数。运用 Mathcad 软件求得

$$a=-2.873\times10^{-4}, b=0.584$$

得到回归曲线

$$r(t)=-2.873\times10^{-4}t+0.584$$

拟合的结果如图 4-1 所示。

我们忽略这种误差，即令$\left|P(t)-r(t)\right|=0$，于是得到我国人口净增长率模型：

$$P(t)=-2.873\times10^{-4}t+0.584 \tag{5.2}$$

我国从 1980 年开始实行计划生育政策。模型（5.2）刻画了两个阶段的人口净增长率，即不仅刻画了没有实行计划生育政策（1950—1979 年)时的人口净增长率又刻画了 1980 年以后的人口净增长率，也就是说反映我国人口净增长率的模型已包含了计划生育政策因素。

5.4.4　模型验证

按照(5.2)式，用数学软件 Mathcad 计算我国 2001—2010 年各年的人口净增长率的预测值，求出平均值，与第六次全国人口普查值 0.0057 比较。

$P(2001)=9.113\times10^{-3}$　　$P(2006)=7.676\times10^{-3}$

$P(2002)=8.825\times10^{-3}$　　$P(2007)=7.389\times10^{-3}$

$P(2003)=8.538\times10^{-3}$　　$P(2008)=7.102\times10^{-3}$

$P(2004)=8.251\times10^{-3}$　　$P(2009)=6.814\times10^{-3}$

$P(2005)=7.963\times10^{-3}$　　$P(2010)=6.527\times10^{-3}$

$\bar{P}=\dfrac{1}{10}\sum_{i=1}^{10}P(200i)=0.0078$，$\left|\bar{P}-0.0057\right|=0.0021$，误差为 2.1‰。

显然，误差比较小，这说明由模型（5.2）能较精确地预测我国未来某个时间段内的人口净增长率。

5.4.5　模型评价

我们在建立模型时是基于年份和人口净增长率具有良好的线性关系这一特点，因而得到适合我国人口增长状况的人口净增长率模型，预测结果比较接近实际值。但是模型还是有一定的误差。事实上，人口净增长率是由生育、死亡、疾病、自然灾害、特殊的人口生育政策、环境、社会以及经济等诸多因素影响和制约的结果，

如此众多的因素不可能通过几个指标表达清楚，它们对人口增长的潜在而复杂的影响更是无法精确计算。这说明人口预测是一个相当复杂的问题。由于数据有限，未能充分考虑影响人口增长的因素，因此所建模型不全面，所得的结果与实际有一定的出入。

习题五

1.在 5.3 节中能否用 4 次、5 次多项式对 33 个离散点进行曲线拟合？

2.在 5.4 节中试用曲线拟合的方法建立我国人口净增长率的数学模型。

3.在 12 小时内，每隔 1 小时测量一次温度，温度依次为：5，8，9，15，25，29，31，30，22，25，27，24（单位：°C）。

①试分别用分段线性插值和三次样条插值方法，估计在 3.2 小时，6.5 小时，7.1 小时，11.7 小时的温度值，每隔 1/10 小时估计一次温度值并画出其图形。

②用多项式拟合，估计在 3.2 小时，6.5 小时，7.1 小时，11.7 小时的温度值。

第 6 章　概率统计建模

在社会、生产、科研和生活实践中，许多问题的不确定现象都是由随机因素的影响所造成的，即将这种现象可以视为一些随机事件，而随机事件一般是按照一定的概率发生的。与此有关的随机因素的变化往往都会服从于一定的概率分布。在实际中，就是利用这些概率分布规律对问题进行研究，从而可以对所研究的实际问题作出估计、推断、预测和决策。实际中的很多问题都是从总体中随机抽取有代表性的一部分（称为样本），通过研究其样本的特性，来统计推断或预测总体的性质。随着社会环境的复杂化和条件的变化，使得日常生活中的不确定问题越来越多，从而概率统计方法在实际中的应用越来越广泛。

6.1　水泥凝固时放出热量问题

6.1.1　问题的提出

某种水泥在凝固时放出的热量 $y(\mathrm{J/g})$ 与水泥中下列 4 种化学成分有关。

x_1 : $3CaO\cdot Al_2O_3$ 的成分（%）

x_2 : $3CaO\cdot SiO_2$ 的成分（%）

x_3 : $4CaO\cdot Al_2O_3\cdot Fe_2O_3$ 的成分（%）

x_4 : $2CaO\cdot SiO_2$ 的成分（%）

现记录了 13 组数据，列在表 6-1 中，根据表中的数据，试研究 y 与 x_1,x_2,x_3,x_4 四种成分的关系。

表 6-1

编号	$x_1(\%)$	$x_2(\%)$	$x_3(\%)$	$x_4(\%)$	y(J / g)
1	7	26	6	60	78.5
2	1	29	15	52	74.5
3	11	56	8	20	104.3
4	11	31	8	47	87.6
5	7	52	6	33	95.9
6	11	55	9	22	109.2
7	3	71	17	6	102.7
8	1	31	22	44	72.5
9	2	54	18	22	93.1
10	21	47	4	26	115.9
11	1	40	23	34	83.8
12	11	66	9	12	113.3
13	10	68	8	12	109.4

6.1.2 模型的建立

为了研究方便，我们考虑一个变量受其他变量影响时，把这个变量称为因变量，记为Y，其他变量称为自变量，记为X，这时相关关系可记作

$$Y = f(x) + \varepsilon \tag{6.1}$$

其中$f(x)$是当$X = x$时，因变量Y的均值，即

$$f(x) = E(Y \mid X = x)$$

称$f(x)$为Y对X的回归函数，ε为Y与$f(x)$的偏差，它是随机变量，并假定$E(\varepsilon) = 0$。

回归函数可以是一元函数，也可以是多元函数，即

$$Y = f(x_1, x_2, \cdots, x_m) + \varepsilon \tag{6.2}$$

其中$f(x_1, x_2, \cdots, x_m) = E(Y \mid X_1 = x_1, X_2 = x_2, \cdots, X_m = x_m)$为$m$元回归函数，统称为多元回归函数。

若回归函数$f(x_1, x_2, \cdots, x_m)$中，$m = 1$且$f(x_1, x_2, \cdots, x_m)$是线性函数，则称$f(x)$为一元线性回归函数；$m > 1$且$f(x_1, x_2, \cdots, x_m)$是多元线性函数，则称其为多元线性回归函数；若回归函数$f(x_1, x_2, \cdots, x_m)$是非线性函数，则称其为非线性回归函数。对非线性回归，经常采用线性化的方法来处理。所以，目前研究最多的是线性回归问题，

且假定 $X_1, X_2, \cdots, X_m$ 和 Y 均服从正态分布。回归分析的任务就是求出满足式（6.2）的回归函数 $f(x_1, x_2, \cdots, x_m)$，从而对所研究的相关关系做出所需的预测和控制。

多元回归模型的应用是相当广泛的，例如，某种商品的销售量可能受收入水平、风俗习惯、产品质量、价格、宣传广告等多种因素的影响；某种产品的质量可能受生产该产品时的温度、湿度、压力、原材料的质量和有害成分的含量等影响；工人的劳动生产率可能受学历、智力水平、情绪的稳定性和才能等因素的影响；某城市的用水量可能与该城市的人口数及工业总产值有关。诸如此类的关系，可以通过多元回归分析模型进行研究。

例如，在水泥凝固时放出热量问题中，可建立线性回归模型

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + \varepsilon \tag{6.3}$$

其中 $E(\varepsilon) = 0$，$D(\varepsilon) = \sigma^2$。而 b_0, b_1, b_2, b_3, b_4 和 σ^2 是未知参数，为了估计这些参数，将表 6-1 的值代入模型（6.3），得线性模型

$$\begin{cases} y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + b_3 x_{i3} + b_4 x_{i4} + \varepsilon_i \\ E(\varepsilon_i) = 0, Cov(\varepsilon_i, \varepsilon_j) = \delta_{ij}\sigma^2, (i, j = 1, \cdots, 13) \end{cases} \tag{6.4}$$

一般地，多元线性回归模型可表示为：

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + \varepsilon \tag{6.5}$$

其中，$x_1, x_2, \cdots x_m$ 是自变量，b_0 为常数，$b_1, b_2, \cdots, b_m$ 为回归系数，$b_0, b_1, b_2, \cdots, b_m$ 皆为未知数，统称 $b_0, b_1, b_2, \cdots, b_m$ 为回归参数，一旦回归参数确定，则多元线性回归模型就完全确定，一般假定随机误差 $\varepsilon \sim N(0, \sigma^2)$。

为了得到回归参数的估计值，就要对变量进行观测，假设对变量的 $n(n > m)$ 次独立观测数据为：$\{(y_i, x_{i1}, x_{i2}, \cdots, x_{im}), i = 1, \cdots, n\}$，则这些观测数据应满足式（6.5），即有

$$\begin{cases} y_1 = b_0 + b_1 x_{11} + b_2 x_{12} + b_3 x_{13} + b_4 x_{14} + \varepsilon_1 \\ y_2 = b_0 + b_1 x_{21} + b_2 x_{22} + b_3 x_{23} + b_4 x_{24} + \varepsilon_2 \\ \cdots\cdots \\ y_n = b_0 + b_1 x_{n1} + b_2 x_{n2} + b_3 x_{n3} + b_4 x_{n4} + \varepsilon_n \end{cases} \tag{6.6}$$

其中 $E(\varepsilon_i) = 0, Cov(\varepsilon_i, \varepsilon_j) = \delta_{ij}\sigma^2, (i, j = 1, \cdots, n)$。

若记 $Y = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}, \beta = (b_0, b_1, \cdots, b_m)^{\mathrm{T}}, \varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)^{\mathrm{T}}$，

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & & x_{2m} \\ \vdots & & & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}_{n\times(m+1)},$$

则多元线性回归的数学模型（6.6）可写成矩阵形式

$$Y = X\beta + \varepsilon \tag{6.7}$$

其中 $E(\varepsilon)=0, Var(\varepsilon)=\sigma^2 I_n$。

（1）参数的最小二乘估计

为了获得参数 β 的估计，我们采用最小二乘法，即选择 β，使

$$Q(\beta)=\sum_{i=1}^{n}\varepsilon_i^2=\varepsilon^{\mathrm{T}}\varepsilon=(Y-X\beta)^{\mathrm{T}}(Y-X\beta) \tag{6.8}$$

达到最小。

将 $Q(\beta)$ 对 β 求导数，并令其为零，得

$$\frac{\partial Q}{\partial \beta}=-2X^{\mathrm{T}}(Y-X\beta)=0$$

即 $X^{\mathrm{T}}X\beta=X^{\mathrm{T}}Y$。记 $L=X^{\mathrm{T}}X$，则

$$L\beta=X^{\mathrm{T}}Y \tag{6.9}$$

方程（6.9）称为正规方程，其中 X 为 $n\times(m+1)$ 阶矩阵，一般假定 $rank(X)=m+1$，由线性代数理论可知 $L=X^{\mathrm{T}}X$ 为满秩矩阵，它的秩 $rank(L)=m+1$，则正规方程（6.9）有唯一解，记作

$$\hat{\beta}=L^{-1}X^{\mathrm{T}}Y \tag{6.10}$$

我们证明（6.10）中 $\hat{\beta}$ 为参数向量 β 的最小二乘法估计量，现用矩阵形式来叙述其证明步骤。

从式（6.8）知，对任意的 β，$Q=(Y-X\beta)^{\mathrm{T}}(Y-X\beta)$，则有

$$\begin{aligned}(Y-X\beta)^{\mathrm{T}}(Y-X\beta)&=[(Y-X\hat{\beta})+X(\hat{\beta}-\beta)]^{\mathrm{T}}[(Y-X\hat{\beta})+X(\hat{\beta}-\beta)]\\&=(Y-X\hat{\beta})^{\mathrm{T}}(Y-X\hat{\beta})+(\hat{\beta}-\beta)^{\mathrm{T}}X^{\mathrm{T}}X(\hat{\beta}-\beta)+(Y-X\hat{\beta})^{\mathrm{T}}X(\hat{\beta}-\beta)\\&\quad+(\hat{\beta}-\beta)^{\mathrm{T}}X^{\mathrm{T}}(Y-X\hat{\beta})\\&=(Y-X\hat{\beta})^{\mathrm{T}}(Y-X\hat{\beta})\end{aligned}$$

注： 上述证明过程中应用了如下结果：

$(\hat{\beta}-\beta)^{\mathrm{T}}X^{\mathrm{T}}X(\hat{\beta}-\beta)=[X(\hat{\beta}-\beta)]^{\mathrm{T}}[X(\hat{\beta}-\beta)]\geq 0$；

$(Y-X\hat{\beta})^{\mathrm{T}}X(\hat{\beta}-\beta)=(Y^{\mathrm{T}}X-\hat{\beta}X^{\mathrm{T}}X)(\hat{\beta}-\beta)=(Y^{T}X-Y^{\mathrm{T}}X)(\hat{\beta}-\beta)=0$。

至此，在 $|L|\neq 0$ 时，证明了式（6.10）中的 $\hat{\beta}$ 是 β 的最小二乘法估计量。

在实际工作中，常称 $\hat{y}=\hat{b}_0+\hat{b}_1x_1+\cdots+\hat{b}_mx_m$ 为经验线性回归方程。

（2）最小二乘法估计量的性质

首先我们在假定 $E(\varepsilon)=0, Var(\varepsilon)=\sigma^2 I_n$ 的条件下，探讨一下由式（6.10）确定 β 的最小二乘法估计量 $\hat{\beta}$ 的性质。

① $\hat{\beta}$ 是 β 的线性无偏估计量。

证明　由于 $\hat{\beta}=L^{-1}X^{T}Y$，每一个 $\hat{b}_i$ 都是 $y_1,\cdots,y_n$ 的线性组合，因而 $\hat{b}_i$ 是 b_i 的线性估计量，此时称 $\hat{\beta}$ 是 β 的线性估计量。

$$\begin{aligned}E(\hat{\beta})&=E(L^{-1}X^{T}Y)=L^{-1}X^{T}E(Y)=L^{-1}X^{T}E(X\beta+\varepsilon)\\&=L^{-1}X^{T}[X\beta+E(\varepsilon)]=L^{-1}X^{T}X\beta=\beta\end{aligned}$$

即 $E(\hat{b}_i)=b_i$， $i=1,\cdots,m$ 。

② $\hat{\beta}$ 的协方差矩阵为 σ^2L^{-1} ，即

$$\begin{cases}D(\hat{b}_i)=\sigma^2c_{ii}\\Cov(\hat{b}_i,\hat{b}_j)=\sigma^2c_{ij},(i,j=0,1,2,\cdots,m+1)\end{cases}$$

其中 $L^{-1}=C=(c_{ij})_{(m+1)\times(m+1)}$ 。

③ $\hat{\beta}$ 是 β 的最小方差线性无偏估计，即在所有线性无偏估计类中，有且只有 $\hat{\beta}$ 使其方差达到最小。

（3）多元线性回归方程的显性检验

从上面的参数估计过程可以看出，对于一批观察数据

$$(y_i,x_{i1},x_{i2},\cdots,x_{im})\qquad i=1,\cdots,n$$

不论它们是否具有线性关系，总可以利用最小二乘法建立起多元线性回归方程

$$\hat{y}=\hat{b}_0+\hat{b}_1x_1+\hat{b}_2x_2+\cdots+\hat{b}_mx_m$$

但是 Y 与 $x_1,x_2,\cdots,x_m$ 是否确实存在相关关系呢？回归方程的效果如何呢？这就要进行"整个回归效果是否显著"的检验。当 $b_1=b_2=\cdots=b_m=0$ 时， y 与 $x_1,x_2,\cdots,x_m$ 没有关系，回归模型没有意义，于是我们要检验 H_0： $b_1=b_2=\cdots=b_m=0$ 是否成立。

若 H_0 成立，则 $x_1,x_2,\cdots,x_m$ 对 y 没有影响；反之，若 H_0 不成立，则 $x_1,x_2,\cdots,x_m$ 对 y 有影响，此时 y 与 $x_1,x_2,\cdots,x_m$ 的线性关系显著，也称为整个回归效果显著。但要注意，即使整个回归效果是显著的，y 也可能只与某几个 x_i 关系密切（相应的 b_i 显著不为零），而与另几个 x_i 关系不密切（相应的 b_i 为零）。这就是说，多元线性回归除了首先要检验"整个回归是否显著"外，还要逐个检验每一个 b_i 是否为零，以便分辨出哪些 x_i 对 y 并无显著影响，最后，还要对各个 b_i 作出区间估计。

为了进行检验和区间估计，可以证明以下结论成立：

① $\frac{1}{\sigma}Q\sim\chi^2(n-m-1)$ ，则 Q 与 $\hat{b}_1,\hat{b}_2,\cdots\hat{b}_m$ 独立。

记 $\overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$ ， $l_{yy}=\sum_{i=1}^{n}(y_i-\overline{y})^2$ ，则称 l_{yy} 为总变差或称为 y 的离差平方和。l_{yy} 可进行如下分解：

$$l_{yy}=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2+\sum_{i=1}^{n}\left(\hat{y}_i-\overline{y}\right)^2=Q+U \tag{6.11}$$

这时 $Q=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 称为残差平方和，$U=\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2$ 称为回归平方和。

记 $s=\sqrt{\dfrac{Q}{n-m-1}}$，称其为剩余标准差或估计的标准差。由于 l_{yy} 不变，当然希望 Q 越小越好，即 U 越大越好，因此，定义复相关系数：

$$R=\sqrt{\frac{U}{l_{yy}}}=\sqrt{1-\frac{Q}{l_{yy}}} \tag{6.12}$$

当观察值 y_i 全都与回归值 $\hat{y}_i$ 吻合时，$Q=0,R=1$；当 $\hat{y}_i=\overline{y}$ 时，$Q=l_{yy},R=0$。在一般情况下，R 的数值在 0 和 1 之间。

复相关系数 R 的定义，类似于两个变量时的相关系数的定义，但要注意，复相关系数 R 只取正值。在两个变量时，有正相关与负相关之分，在多个变量时，就没有这一说法了，所以复相关系数 R 只取正值。

②在 $b_1=b_2=\cdots=b_m=0$ 的条件下，

$$\frac{U}{\sigma^2}\sim\chi^2(m)$$

且 U 与 Q 独立，因此

$$F=\frac{U/m}{Q/(n-m-1)}=\frac{n-m-1}{m}\frac{R^2}{1-R^2}\sim F(m,n-m-1)$$

③$t_i=\dfrac{\hat{b}_i-b_i}{\sqrt{c_{ii}}\cdot s}\sim t(n-m-1)$，此时，有

$$F_i=\frac{(\hat{b}_i-b_i)^2/c_{ii}}{Q/(n-m-1)}=\frac{(\hat{b}_i-b_i)^2}{c_{ii}^2s^2}\sim F(1,n-m-1),\ i=1,2,\cdots,m$$

这里，c_{ii} 为 $C=L^{-1}$ 中第 i 个对角线元素。

利用上述几条结论，可进行下列检验、估计和预测。

①回归显著性检验（F 检验）

该检验是考察整个回归效果是否显著的。若整个回归效果不显著，即全部回归系数为零。因此，设原假设 H_0：$b_1=b_2=\cdots=b_m=0$。若 H_0 为真，则

$$F=\frac{n-m-1}{m}\frac{U}{Q}=\frac{n-m-1}{m}\cdot\frac{R^2}{1-R^2}\sim F(m,n-m-1)$$

而且在 H_0 不成立时，F 值有变大的趋势，因此应取右侧否定域，故检验法是当 $F>F_\alpha(m,n-m-1)$ 时拒绝原假设，认为回归效果显著；否则认为回归效果不显著。

②单个回归系数为零的检验（t 检验）

该检验即某个自变量是否对因变量有显著性影响的检验。

在多元回归分析中可能出现 y 与所有自变量的总体是有相关关系的，但 y 与某个特定的 x_i 则可能无关，即 x_i 对 y 并不起作用或者已被其他的 x_i 的作用所代替，为此设 m 个原假设 $H_{0i}: b_i = 0$，$i = 1,2,\cdots,m$。

若 H_{0i} 为真，则统计量 $t_i = \dfrac{\hat{b}}{\sqrt{c_{ii}\cdot s}} \sim t(n-m-1), i = 1, 2, \cdots, m$。而当 H_{0i} 不成立时，$|t_i|$ 有变大的趋势，因而应取双侧拒绝域。故当 $|t_i| > t_{\frac{\alpha}{2}}(n-m-1)$ 时，否定 H_{0i}，即认为 x_i 对 y 是有作用的；若某几个 x_i 是有作用的，而另几个 x_i 是不起作用的，则应从回归方程中删除那些不起作用的自变量。

单个回归系数是否为零，也可以用 F 检验，即若 H_{0i} 为真，统计量

$$F_i = \frac{\hat{b}_i^2}{c_{ii}^2 s^2} \sim F(1, n-m-1), i = 1, 2, \cdots, m$$

故当 $F_i > F_\alpha(1,\ n-m-1)$ 时，拒绝原假设，即认为 x_i 对 y 的影响是显著的；否则认为 x_i 对 y 的影响是不显著的。

③对 b_i 的区间估计

由于 $\dfrac{\hat{b}_i - b_i}{\sqrt{c_{ii}}\cdot s} \sim t(n-m-1)$，因而 b_i 的 $1-\alpha$ 置信区间为 $(\hat{b}_i - d_i,\ \hat{b}_i + d_i)$，其中 $d_i = t_{\frac{\alpha}{2}}(n-m-1)\cdot\sqrt{c_{ii}}\cdot s$。

④ y_0 的 95%预测区间近似为 $(\hat{y}_0 - 2s,\ \hat{y}_0 + 2s)$，其中

$$\hat{y}_0 = \hat{b}_0 + \hat{b}_1 x_{01} + \hat{b}_2 x_{02} + \cdots + \hat{b}_m x_{0m}$$

（4）多元线性回归分析模型的推广

①多项式回归分析模型

类似于模型（6.5），由自变量多项式的随机项组成的回归模型称为多项式回归模型，它的一般形式为：

$$Y = b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m + \varepsilon \qquad (6.13)$$

初看模型（6.13）不是线性回归，因自变量中含有幂函数，但由于未知参数 $x^i, i = 1,\cdots,m$ 都是线性出现的，因此令

$$x_1 = x, x_2 = x^2, \cdots, x_m = x^m$$

则模型（6.13）就变为多元线性回归模型：

$$Y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_m x_m + \varepsilon$$

从而多项式回归模型可以用多元线性回归模型的计算公式和检验方法。多项式回归还有许多推广的形式，例如：

A. $y = b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m + \dfrac{c}{x}$；

B. $y = b_0 + b_1 x \ + b_2 x^2 + \cdots + b_m x^m + c \ln x$；

C. $y = \exp(b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m)$；

D. $y = \exp(b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m + \dfrac{c}{x})$；

E. $y = \exp(b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m) x^c$。

②广义线性回归模型

广义线性回归模型的一般形式为：

$$y = f\left[b_0 + b_1 F_1(x_1, x_2, \cdots, x_m) + \cdots + b_p F_p(x_1, x_2, \cdots, x_m)\right]$$

其中，$y = f(y_0)$是一个不含未知参数的一元函数，有反函数$y_0 = g(y)$。$F_j = (x_1, x_2, \cdots, x_m)\,(j = 1, 2, \cdots, p)$是$x_1, x_2, \cdots, x_m$的不含未知参数的多元函数。

广义线性回归模型的回归系数的确定主要是从自变量$x_1, x_2, \cdots, x_m$和因变量y以及n组观察值$(x_{i1}, x_{i2}, \cdots, x_{im}, y_i)$, $i = 1, 2, \cdots, n$出发，用最小二乘法求出$b_0, b_1, \cdots, b_p$的估计$\hat{b}_0, \hat{b}_1, \cdots, \hat{b}_p$，使得

$$Q = \sum_{i=1}^{n}[g(y_i) - (b_0 + b_1 F_1(x_{i1}, x_{i2}, \cdots, x_{im}) + \cdots + b_p F_p(x_{i1}, x_{i2}, \cdots, x_{im}))]^2$$

达到最小。此时令

$$\begin{cases} Y = g_0(y) \\ t_1 = F_1(x_1, x_2, \cdots, x_m) \\ \cdots \\ t_p = F_p(x_1, x_2, \cdots, x_m) \end{cases}$$

则$Y = b_0 + b_1 t_1 + \cdots + b_p t_p$，这样就把广义线性回归模型化为多元线性回归模型。

（5）MATLAB 和 Mathematica 求解

①MATLAB 命令

命令格式 [b，bint，r，rint，stats]=regress(Y，X，alpha)，其中输入向量 X，Y 的排列方式分别为

$$X = \begin{bmatrix} 1, x_{11}, x_{22}, \cdots, x_{1m} \\ 1, x_{21}, x_{22}, \cdots, x_{2m} \\ \vdots \\ 1, x_{n1}, x_{n2}, \cdots, x_{nm} \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

alpha 为显著性水平（缺省时设定为 0.05）。

输出向量 b 为回归系数的估计值，即$b = (b_0, b_1, \cdots, b_n)^{\mathrm{T}}$；

输出向量 bint 为回归系数估计值的置信区间；

输出向量 r 为残差向量；

输出向量 rint 为残差向量的置信区间；

输出向量 stats $=(R^2, F, P)^T$，它是一个 3 维向量，用于检验回归模型的统计量，其中第一个分量 R^2 中的 R 是相关系数，第二个分量是 F 统计量，第三个分量是与统计量 F 对应的概率 P，当 $P <$ alpha 时拒绝原假设 H_0，说明回归模型成立。

②Mathematica 命令

Mathematica 中键入命令<<Statisti\Linearre.m，按 Shift + Enter 键，即可调入线性回归软件包。输入

$$data=\{\{x_{11},x_{12},\cdots,x_{1m},y_1\},\{x_{21},x_{22},\cdots,x_{2m},y_2\},\cdots,\{x_{n1},x_{n2},\cdots,x_{nm},y_n\}\};$$

$$Regress=[data,\{1,x_1,x_2,\cdots,x_m\},\{x_1,x_2,\cdots,x_m\},OutptList\to\{BestFit\}]$$

（6）实际问题的求解

水泥凝固时放出热量问题

在 MATLAB 编辑器中输入以下程序：

```
%水泥放出热量问题 ch411
  %  文件名：ch411.m
    x1 =[7,1,11,11,7,11,3,1,2,21,1,11,10]';
    x2 =[26,29,56,31,52,55,71,31,54,47,40,66,68]';
    x3 =[6,15,8,8,6,9,17,22,18,4,23,9,8]';
    x4 =[60,52,20,47,33,22,6,44,22,26,34,12,12]'
    y =[78.5,74.3,104.3,87.6,95.9,109.2,102.7,72.5,93.1,115.9,83.8,113.3,109.4];
    x =[ones(13,1),x1,x2,x3,x4]
[b，bint，r，rint，stats] = regress(y，x，0.05);
    disp (‘回归系数估计值’)
    b
    disp(‘回归系数估计值的置信区间’)
    bint
    disp(‘残差平方和’)
    r'*r
    disp( ‘相关系数的平方’)
    stats(1)
```

```
disp( ‘F 统计量’ )
stats(2)
disp( ‘与统计量 F 对应的概率 p’ )
stats(3)
```

执行后输出

回归系数估计值

```
b =
62.4054
1.5511
0.5102
0.1019
-0.1441
```

回归系数估计值的置信区间

```
bint =
-99.1786    223.9893
-0.1663       3.2685
-1.1589       2.1792
-1.6385       1.8423
-1.7791       1.4910
```

残差平方和

```
ans =
47.8636
```

相关系数的平方

```
ans =
0.9824
```

F 统计量

```
ans =
111.4792
```

与统计量 *F* 对应的概率 *P*

```
ans =
4.7562e – 007
```

从计算结果可知，回归方程

$$y = 62.405 + 1.551x_1 + 0.5102x_2 + 0.1019x_3 - 0.1441x_4$$

查表得：

$$F_{0.05}(m, n-m-1) = F_{0.05}(4.8) = 3.838$$

易见统计量 $F = 111.4792 > F_{0.05}(4.8) = 3.838$。进一步可得 $F > t_{0.05}(4.8) = 7.006$，所以回归效果是高度显著的。

6.2　决策模型

6.2.1　问题的提出

决策是人们在生活和工作中普遍存在的一种活动，是为解决当前或未来可能发生的问题，选择最佳方案的一种过程。比如，某人要到某地出差，而天气预报可说能有寒流，考虑出差是否要带棉大衣，带上棉大衣无寒流是个累赘，若不带又可能遇上寒流而挨冻，到底带不带？这就要他作出决策。又如生产某种产品的工厂，若对此种产品的市场需求不是很了解，生产的数量太小，影响企业收入，生产的数量太大，又势必造成产品积压，影响资金周转，给企业造成损失，到底生产多少为宜？这就需要有关人员通过市场调查后作出决策。所以，小到个人生活，大至企业经营以及国家的政治经济问题，都需要决策。本节介绍决策的一些基本术语和常见的两种决策方法。

例 1　某公司为了扩大市场，要举办一个产品展销会，会址打算选择甲、乙、丙三地；获利情况除了与会址有关外，还与天气有关，天气分为晴、阴、雨三种，据气象台预报，估计三种天气情况可能出现的概率分别为 0.20，0.50，0.30，其收益情况如表 6－2，现要通过分析确定会址，使收益最大。

表　6-2

选址 \ 效益 \ 天气 / 概率	N_1(晴)	N_2(阴)	N_3(雨)
	P_1=0.20	P_2=0.50	P_3=0.30
A_1(甲址)	4	6	1
A_2(乙址)	5	4	1.5
A_3(丙址)	6	2	1.2

6.2.2 决策的概念和类型

在决策问题中，把面临的几种自然情况叫自然状态或客观条件，简称状态或条件。如例 1 中的 N_1, N_2, N_3 就是各种不同的自然状态，这些是不可控因素，但只能有一种出现。把 A_1, A_2, A_3 称为行动方案或策略，这些是可控因素，由决策者决定。表 6-2 中后三行数字称为益损值，根据这些数字的含义，有时也叫效益值或损失值，由它们构成的矩阵

$$M = \begin{bmatrix} 4 & 6 & 1 \\ 5 & 4 & 1.5 \\ 6 & 2 & 1.2 \end{bmatrix}$$

叫做决策的益损矩阵或风险矩阵。P_1, P_2, P_3 是各状态出现的概率。

一般地，如决策问题的可控因素（即行动方案）用 $A_i (i=1,2,\cdots,m)$ 表示，状态用 $N_j (j=1,2,\cdots,n)$ 表示，在 N_j 状态下采用 A_i 行动方案的风险值用 a_{ij} 表示，N_j 状态出现的概率用 P_j 表示，则可根据 n 的大小和 P_j 的信息情况，将决策问题分为三类：确定型决策、风险型决策和不确定型决策。

当 n=1 时，决策问题就是确定型的，我们主要讨论风险型和不确定型的决策问题。

1.风险型决策

当 $n>1$，且各种自然状态出现的概率 $P_j (i=1,2,\cdots,n)$ 可通过某种途径获得时，决策问题就是风险型决策问题。如例 1 就是风险型决策问题，对于这类问题，我们介绍两种决策准则和相应的解决方法。

（1）最大可能准则

由概率论知识，一个事件的概率就是该事件在一次试验中发生的可能性大小，概率越大，事件发生的可能性就越大。基于这种思想，在风险决策中我们选择一种发生概率最大的自然状态来进行决策，而不顾及其他自然状态的决策方法，这就是最大可能准则。这个准则的实质是将风险型决策问题转化为确定型决策问题的一种决策方法。

若对例 1 按最大可能准则进行决策，则因为自然状态 N_2 出现的概率 $P_2 = 0.50$ 最大，因此就在这种自然状态下进行决策，通过比较可知，采取 A_1 行动方案获利最大。因此，采用 A_1 方案是最优决策。

注 如果各自然状态的概率较接近，一般不使用这种决策准则。

（2）期望值准则（决策树法）

如果把每个行动方案看作随机变量，每个自然状态下的效益值看作随机变量的

取值，其概率为自然状态出现的概率，则期望值准则就是将每个行动方案的数学期望计算出来，视其决策目标的情况选择最优行动方案。

若对例 1 按期望值准则进行决策，则需要计算各行动方案的期望收益，事实上

$$E(A_1)=4\times0.2+6\times0.5+1\times0.3=4.1$$
$$E(A_2)=5\times0.2+4\times0.5+1.5\times0.3=3.45$$
$$E(A_3)=6\times0.2+2\times0.5+1.2\times0.3=2.56$$

显然，$E(A_1)$ 最大，所以采取行动方案 A_1 最佳，即选择甲地举办展销会效益最大。

值得注意的是，为了形象直观地反映决策问题未来发展的可能性和可能结果所作的预测而采用的决策树法就是按期望值准则进行决策的一种方案。以例 1 来说明其决策步骤。

例 1 的决策树如图 6－1 所示，其中：

□表示决策点，从它引出的分枝叫方案分枝，其数目就是方案数。

○表示机会节点，从它引出的分枝叫概率分枝，每条概率分枝代表一种自然状态，并标有相应状态发生的概率。

△称为末梢节点，右边数字表示各方案在不同自然状态下的益损值。

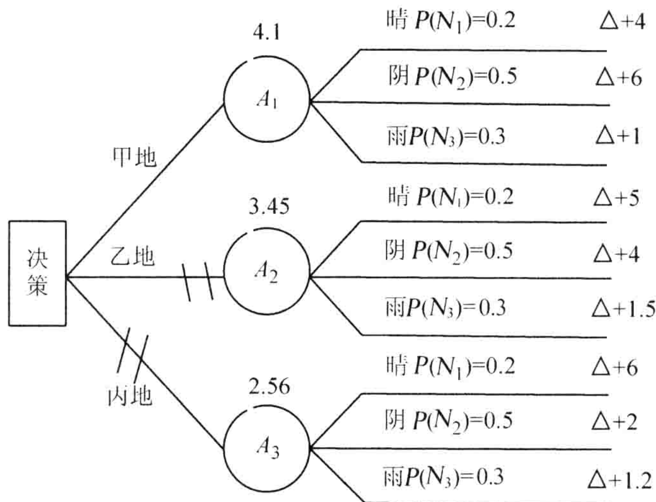


图 6-1　决策树

计算各机会节点的期望值，并将结果标在节点上方，再比较各机会节点上标值的大小，进行决策，在淘汰方案分枝上标“//”号，余下方案即为最优方案，最优方案的期望值标在决策点的上方。由于 A_1 上方标值 4.1 为最大，因此选定方案 A_1，其收益数值的期望为 4.1。

此例只包括一个决策点，称为单级决策问题。在有些决策问题中至少包括两个

决策点，称为多级决策问题，可利用同样的思路进行决策。

例 2 某工程采用正常速度施工，若无坏天气的影响，可确保在 30 天内按期完成工程，但据天气预报，15 天后天气肯定变坏，有 40% 的可能出现阴雨天气，但这不会影响工程进度，有 50% 的可能遇到小风暴，而使工期延长 15 天；另有 10% 的可能遇到大风暴而使工期延长 20 天。对于以上可能出现的情况，考虑两种方案：

（1）提前加班，确保工程在 15 天内完成，实施此方案需增加额外支付 18 000 元。

（2）先维持原定的施工进度，等到 15 天后根据实际出现的天气状况再作对策：

①若遇阴雨天，则维持正常进度，不必支付额外费用。

②若遇小风暴，则有下述两个供选方案：

一是抽空（风暴过后）施工，支付工程延期损失费 20000 元。

二是采用应急措施，实施此措施可能有三种结果：有 50% 的可能减少误工期 1 天，支付延期损失费和应急费用 24000 元；30% 的可能减少误工期 2 天，支付延期损失费和应急费用 18000 元；有 20% 的可能减少误工期 3 天，支付延期损失费和应急费用 12000 元。

③若遇大风暴，则仍然有两个方案可供选择：

一是抽空进行施工，支付工程的延期损失费 50000 元；

二是采取应急措施，实施此措施可能有三种结果：有 70% 的可能减少误工期 2 天，支付延期损失费及应急费用 54000 元；有 20% 可能减小误工期 3 天，支付延期损失费及应急费用 46000 元；有 10% 的可能减少误工期 4 天，支付延期损失费及应急费用 38000 元。试进行决策，选择最佳行动方案。

解 （1）据题意画出决策树，如图 6－2。

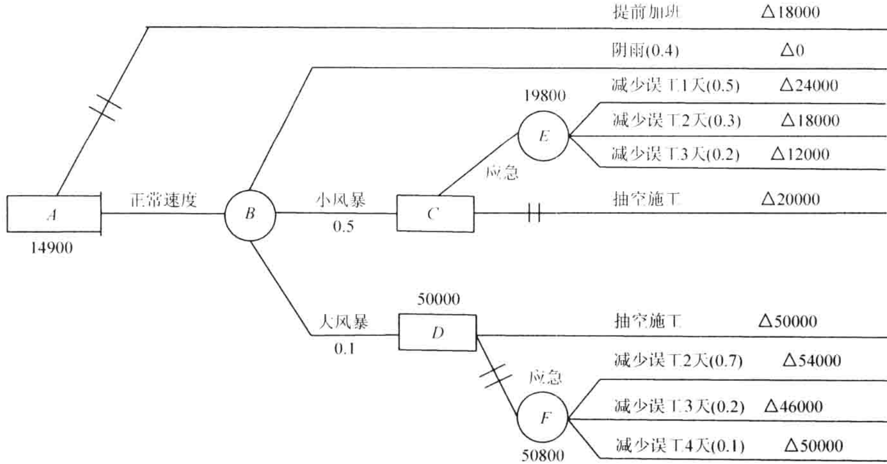


图 6-2 决策树

（2）计算第一级机会点 E，F 的损失费用期望值

$$E(E) = 0.5\times 24000 + 0.3\times 18000 + 0.2\times 12000 = 19800$$

$$E(F) = 0.7\times 54000 + 0.2\times 46000 + 0.1\times 38000 = 50800$$

将 19800 和 50000 标在相应的机会点上，然后在第一级决策点 C，D 外分别进行方案比较：首先考察 C 点，其应急措施支付额外费用的期望值较少，故它为最佳方案，同时划去抽空施工的方案分枝，再在 C 上方标明最佳方案期望损失费用 19800 元；再考虑 D 处的情况，应急措施比抽空施工支付的额外费用的期望值少，故划去应急措施分枝，在 D 上方标上 50000 元。

（3）计算第二级机会点 B 的损失费用期望值

$$E(B) = 0.4\times 0 + 0.5\times 19800 + 0.1\times 50000 = 14900$$

将其标在 B 的上方，在第二级决策点 A 处进行比较，发现正常进度方案为最佳方案，故划去提前加班的方案分枝，并将 14900 标在 A 点下方。

因此，合理的决策应是开始以正常施工进度进行施工，15 天后再根据具体情况作进一步决策，若出现阴雨天，则维持正常速度；若出现小风暴可采用应急措施；若出现大风暴，则进行抽空施工。

2.不确定型决策

风险型决策问题的自然状态发生的概率既不知道、也无法预先估计或利用历史资料得到时的决策问题就称为不确定型决策问题。仍用 $N_1, N_2, \cdots, N_n$ 表示决策问题中的自然状态，$A_1, A_2, \cdots, A_m$ 表示行动方案，a_{ij} 表示在自然状态 N_j 下采用 i 种行动方案的益损值。若 a_{ij} 为效益值时取正值；若 a_{ij} 为损失值时取负值。

下面介绍几种不确定型的决策准则。

（1）乐观准则

乐观准则的思想就是对客观情况总是持乐观态度，事事都合人意，即选最大效益的最大值 $\max_i\left\{\max_j a_{ij}\right\}$ 所对应的行动方案作为决策。

（2）悲观准则

悲观准则的思想就是对客观情况总是持悲观态度，万事都不会如意，即总是把事情的结果估计得很不利，因此就在最坏的情况下找一个较好的行动方案。也就是在每个状态下的最小效益值中选最大值 $\max_i\{\min_j a_{ij}\}$ 所对应的行动方案作为决策。

（3）等可能准则

等可能准则的思想就是既然不能断定哪种自然状态出现的可能性的大小，就认为各自然状态出现的可能性相同，即 $P(N_j) = \dfrac{1}{n}, j = 1, 2, \cdots, n$。然后按风险型决策的

方法进行决策。

例 3 某厂有一种新产品，其推销策略有 S_1, S_2, S_3 三种可供选择，但各方案所需资金、时间都不同，加上市场情况的差别，因而获利和亏损情况不同，而市场情况有三种可能：N_1 需求量大，N_2 需求量一般，N_3 需求量低。市场情况的概率并不知道，其效益值见表 6－3。

表 6-3

a \ N / A		市场情况		
		N_1	N_2	N_3
销售策略	S_1	50	10	-5
	S_2	30	25	0
	S_3	10	10	10

试用：

（1）乐观法进行决策；

（2）悲观法进行决策；

（3）等可能法进行决策。

解 （1）因为每个行动方案在各种状态下的最大效益值为

$$\max_j \{a_{1j}\} = \max\{50, 10, -5\} = 50$$

$$\max_j \{a_{2j}\} = \max\{30, 25, 0\} = 30$$

$$\max_j \{a_{3j}\} = \max\{10, 10, 10\} = 10$$

所以最大效益的最大值为

$$\max_i \left\{\max_j a_{ij}\right\} = \max\{50, 30, 10\} = 50$$

其最大值 50 对应的行动方案为 S_1，因此用乐观法的决策结果是执行策略 S_1。

（2）因为每个行动方案在各种状态下的最小效益值为

$$\min_j \{a_{1j}\} = \min\{50, 10, -5\} = -5$$

$$\min_j \{a_{2j}\} = \min\{30, 25, 0\} = 0$$

$$\min_j \{a_{3j}\} = \min\{10, 10, 10\} = 10$$

所以，最小效益值的最大值为

$$\max_{i}\left\{\min_{j} a_{ij}\right\}=\max\{-5,0,10\}=10$$

其最大值 10 对应的行动方案为 S_3。因此用悲观法决策的结果是应执行策略 S_3。

（3）取 $p(N_j)=\dfrac{1}{3}, j=1,2,3;$ 计算出各行动方案的期望值为

$$E(S_1)=\frac{1}{3}\times 50+\frac{1}{3}\times 10+\frac{1}{3}\times(-5)=\frac{55}{3}$$

$$E(S_2)=\frac{1}{3}\times 30+\frac{1}{3}\times 25+\frac{1}{3}\times 0=\frac{55}{3}$$

$$E(S_3)=\frac{1}{3}\times 10+\frac{1}{3}\times 10+\frac{1}{3}\times 10=10$$

显然 $E(S_1)=E(S_2)$ 都达到最大值，这时究竟选哪一个策略可由决策者的偏好决定，若是乐观型的，可选 S_1，否则选 S_2。

6.3 最佳订票问题

6.3.1 问题的提出

在激烈的市场竞争中，航空公司为争取更多的客源而开展的一个优质服务项目是预订票业务。公司承诺，预先订购机票的乘客如果未能按时前来登机，可以乘坐下一班飞机或退票，无需附加任何费用。当然也可以订票时只订座，登机时才付款，这两种办法对于下面的讨论是等价的。

设某种型号的飞机容量为 n，若公司限制预订 n 张机票，那么，由于总会有一些订了机票的乘客不按时来登机，致使飞机因不满员飞行而利润降低，甚至亏本，如果不限制订票数量呢，那么当持票按时前来登机的乘客超过飞机容量时，必然会引起那些不能登机飞走的乘客（以下称被挤掉者）的抱怨。公司不管以什么方式予以补救，都会导致受到一定的经济损失，如客源减少，或挤到以后班机的乘客，公司要无偿供应食宿或者付给一定的赔偿金等。这样，综合考虑公司的经济利益，必然存在一个恰当的订票数量和限额。

假设飞机容量为 300，乘客准时到达机场而未乘上飞机的赔偿费是机票价格的 10%，飞行费用与飞机容量、机票价格成正比（由统计资料知，比例系数为 0.6，乘客不按时前来登机的概率为 0.03），请你：

（1）建立一个数学模型，给出衡量公司经济利益和社会声誉的指标，对上述预订票业务确定最佳的预订票数量。

（2）考虑不同客源的不同需要，如商人喜欢上述这种无约束的预订票业务，他们宁愿接受较高的票价；而按时上下班的雇员或游客，愿意以若不能按时前来登机，

则机票失效为代价，换取较低额的票价。公司为降低风险，可以把后者作为基本客源。根据这种实际情况，制定更好的预订票策略。

6.3.2 模型假设与符号说明

1.模型假设

（1）假设预订票的乘客是否按时前来登机是随机的。

（2）假设已预订票而不能前来登机的乘客数是一个随机变量。

（3）假设飞机的飞行费用与乘客的多少无关。

2.符号说明

n：飞机座位数，即飞机的容量；

g：机票的价格；

f：飞行费用；

b：乘客准时到达机场而未乘上飞机的赔偿费；

m：售出机票数；

k： 已预订票而不能前来登机的乘客数，即迟到的乘客数，它是一个随机变量；

p_k：已预订票的 m 个乘客中有 k 个乘客不能按时前来登机的概率；

p：每位乘客迟到的概率；

$p_j(m)$：已预订票前来登机的乘客中至少挤掉 j 人的概率，即社会声誉指标；

S：公司的利润；

ES：公司的平均利润。

6.3.3 模型建立

通过上面引进的符号易知，赔偿费 $b=0.1g$，飞行费用 $f=0.6ng$，每位乘客迟到的概率 $p=0.03$，已预订票的 m 个乘客中，恰有 k 个乘客不能按时前来登机，即迟到的乘客数 k 服从二项分布 $B(m,p)$，此时，

$$p_k=\mathrm{C}_m^k p^k(1-p)^{m-k}\qquad (k=0,1,2,\cdots,m)$$

当 $m-k\le n$ 时，说明 $m-k$ 个乘客全部登机，此时利润

$$S=(m-k)g-f$$

当 $m-k>n$ 时，说明有 n 个乘客登机，有 $m-k-n$ 个乘客没有登上飞机，即被挤掉了，此时利润

$$S=ng-f-(m-k-n)b$$

根据以上分析，利润 S 可表示为：

$$S=\begin{cases}(m-k)g-f & m-k\le n\ \ (k\ge m-n)\\ ng-f-(m-k-m)b & m-k>n\ \ (k<m-n)\end{cases}$$

迟到的乘客数$k=0,1,2,\cdots,m-n-1$时，说明有$m-k-n$个乘客被挤掉了；迟到的乘客数$k=m-n,m-n+1,\cdots,m$时，说明已来的$m-k$个乘客全部登机了。于是平均利润

$$ES=\sum_{k=0}^{m-n-1}(ng-f-(m-k-n)b)p_k+\sum_{k=m-n}^{m}[(m-k)g-f]p_k$$

因为

$$\sum_{k=m-n}^{m}[(m-k)g-f]p_k=(mg-f)(1-\sum_{k=0}^{m-n-1}p_k)-g(\sum_{k=0}^{m}kp_k-\sum_{k=0}^{m-n-1}kp_k)$$
$$=(mg-f)-(mg-f)\sum_{k=0}^{m-n-1}p_k-gE(k)+g\sum_{k=0}^{m-n-1}kp_k$$

所以

$$ES=\sum_{k=0}^{m-n-1}(ng-f-(m-k-n)b)p_k+(mg-f)-(mg-f)\sum_{k=0}^{m-n-1}p_k-gE(k)+g\sum_{k=0}^{m-n-1}kp_k$$
$$=(m-E(k))g-f+\sum_{k=0}^{m-n-1}(ng-f-(m-k-n)b-(mg-f)+gk)p_k$$
$$=(m-E(k))g-f-(b+g)\sum_{k=0}^{m-n-1}(m-k-n)p_k$$

由于$k\sim B(n,p)$，$p_k=\mathrm{C}_m^k p^k(1-p)^{m-k}$，因此，随机变量$k$的数学期望$E(k)=mp$，此时，

$$ES=(1-p)mg-f-(b+g)\sum_{k=0}^{m-n-1}(m-k-n)\mathrm{C}_m^k p^k(1-p)^{m-k}$$

通过以上对问题的分析，可以在一定的社会声誉指标$P_j(m)$范围内，寻求合适的m，根据$f=0.6ng$的关系，使得目标函数ES/f达到最大，即

$$\max\frac{ES}{f}=\frac{1}{0.6N}[(1-p)m-(1+\frac{b}{g})\sum_{k=0}^{m-n-1}(m-k-n)\mathrm{C}_m^k p^k(1-p)^{m-k}]-1$$
$$=\frac{1}{180}[0.97m=1.1\sum_{k=0}^{m-n-1}(m-k-300)\mathrm{C}_m^k p^k(1-p)^{m-k}]-1$$

下面考虑社会声誉指标。

由于$m=n+k+j$，所以$k=m-n-j$，即当被挤掉的乘客数为j时，等价的说法是恰有$m-n-j$个迟到的乘客。

公司希望被挤掉的乘客人数不要太多，被挤掉的概率不要太大，可用至少挤掉j人的概率作为声誉指标，相应地k的取值范围为$k=0,1,2,\cdots,m-n-j$，社会声誉指标

$$P_j(m)=\sum_{k=0}^{m-n-j}\mathrm{C}_m^k p^k(1-p)^{m-k} \tag{6.14}$$

6.3.4 模型求解

为了对模型（6.14）进行求解，可以分别给定m，比如$m=305,306,\cdots,350$，

计算 ES/f；同时，给定 j，比如取 $j=5$，计算社会声誉指标 $P_j(m)$，从中选取使 ES/f 最大，且社会声誉指标 $P_j(m)$ 小于等于某个 α（比如取 $\alpha=0.05$）的最佳订票数 m。

下面给出 MATLAB 计算程序。

```
%飞机最佳订票策略 ch43
%文件名：ch43.m
%m 表示售出的票数；Es 表示平均利润；p 表示声誉指标；
for m=305：325
sm=0；
p=0；
for k=0；m-305
pp=(prod(m-k+1:m)/prod(1:k))*0.03^k*0.97^(m-k)；
p=p+pp；
sm=sm+(m-k-300)*pp/prod(1:k)；
end
Es=(1/180)*[0.97*m-1.1*sm]-1；
M
Es
p
end
```

执行后输出以下结果：

m	ES	P
305	0.6436	9.2338e-005
306	0.6490	9.3723e-004
307	0.6543	0.0048
308	0.6596	0.0167
309	0.6649	0.0442
310	0.6703	0.0952
311	0.6756	0.1742
312	0.6810	0.2796
313	0.6864	0.4028

314	0.6917	0.5314
315	0.6971	0.6525
316	0.7024	0.7566
317	0.7078	0.8388
318	0.7132	0.8890
319	0.7185	0.9399
320	0.7239	0.9661
321	0.7293	0.9818
322	0.7347	0.9907
323	0.7400	0.9954
324	0.7454	0.9979
325	0.7508	0.9990

从计算结果易见，当 m=309 时，社会声誉指标 $p_s(309)=0.0442<0.05$, 当 m=310 时，社会声誉指标 $p_s(310)=0.0952>0.05$, 所以为了使 ES/f 尽量大，且要满足社会声誉指标 $p_s(m)<0.05$, 则最佳订票数可取 m=309。

6.3.5 问题进一步讨论

对于问题的第二个小问题，可将乘客分为两类：一类如商人可采用较高的票价；另一类乘客，比如上下班的雇员或游客，可设 m 张预订票中有 ω 张折扣票，折扣票价为 $\alpha g(\alpha<1)$，但若不按时前来登机，责任自负。然后做出合理的简化假设，可建立类似的模型，计算并分析结果。注意这里多了两个参数 α 和 ω。

6.4 存储模型

本节我们讨论含有随机数据的存储模型。为此，我们先通过一个例题介绍建立这种模型的基本思想。

例 1　某商店拟在新年期间出售一批挂历。根据成本核算，每售出一千本可获利 700 元。如果在新年期间售不完，新年后必须削价处理，这样每一千本就要亏损 400 元。根据往年的经验，市场需求的概率如表 6-4。

若每年只能订货一次，问应订购挂历几千本，才能使获利的期望值最大？

表 6-4

需求量 r（千本）	0	1	2	3	4	5
概率 $P(r)$	0.05	0.10	0.25	0.35	0.15	0.10

解 假设该商店订货 4000 本，计算一下获利情况。

当市场需求为 0 千本时，获利为 $-400\times4=-1600$（元）

当市场需求为 1 千本时，获利为 $-400\times3+700=-500$（元）

当市场需求为 2 千本时，获利为 $-400\times2+700\times2=600$（元）

当市场需求为 3 千本时，获利为 $-400\times1+700\times3=1700$（元）

当市场需求为 4 千本时，获利为 $700\times4=2800$（元）

当市场需求为 5 千本时，获利为 $700\times4=2800$（元）

可见，订货量为 4 千本时获利的期望值为

$$E\left(C(4)\right)=-1600\times0.05+(-500)\times0.10+600\times0.25+1700\times0.35+2800\times0.15+2800\times0.10=1315\text{ 元}$$

按上述计算法，可得表 6-5。

表 6-5

获利 需求量 定购量	0	1	2	3	4	5	获利的期望值(元)
0	0	0	0	0	0	0	0
1	-400	700	700	700	700	700	645
2	-800	300	1400	1400	1400	1400	1180
3	-1200	-100	1000	2100	2100	2100	1440*
4	-1600	-500	600	1700	2800	2800	1315
5	-2000	-900	200	1300	2400	3500	1025

由此表可知，获利最大期望值为 1440 元（标*者），即该商店订购 3000 本挂历可使获利期望值最大。

注 本例也可以从相反的角度求解。当订货量为 Q 时，可能发生供过于求或供不应求的情况，把这两种损失合起来考虑，取损失期望值最小者为所对应的 Q 值。

6.4.1 需求是随机离散变量的存储策略模型一

报童问题 报童每天的售报数量是一个随机变量。假设报童每售出一份报纸可获利 k 元，而未能售出的报纸，每份要赔 h 元，根据以往经验，报童每日售出报纸

数量 r 的概率 $P(r)$ 是已知的。报童每日应订购多少份报纸才能获利最多？

我们用计算损失期望值最小的方法建立模型。设报童每日订购报纸的数量为 Q。

（1）供过于求 $(r \le Q)$ 时，因不能售出而承担的损失期望值为

$$\sum_{r=0}^{Q} h(Q-r)P(r)$$

（2）供不应求 $(r > Q)$ 时，因缺货而少挣钱的损失期望值为

$$\sum_{r=Q+1}^{\infty} k(r-Q)P(r)$$

于是所求模型为

$$\min C(Q) = h\sum_{r=0}^{Q}(Q-r)P(r) + k\sum_{r=Q+1}^{\infty}(r-Q)P(r)$$

由于 Q 是离散的，故不能直接用微积分方法求解。为解此模型，我们用差分近似微分，设 $\Delta C(Q) = C(Q+1) - C(Q)$ 为 $C(Q)$ 的差分。

由于

$$C(Q+1) = h\sum_{r=0}^{Q+1}(Q+1-r)P(r) + k\sum_{r=Q+2}^{\infty}(r-Q-1)P(r)$$

$$C(Q) = h\sum_{r=0}^{Q}(Q-r)P(r) + k\sum_{r=Q+1}^{\infty}(r-Q)P(r)$$

所以

$$\begin{aligned}\Delta C(Q) &= C(Q+1) - C(Q) \\ &= (h+k)\sum_{r=0}^{Q} P(r) - k\end{aligned}$$

令 $\Delta C(Q) = 0$，即订货量多一张和少一张时，损失的期望值相同，由此可近似地求出驻点 Q（使 $C(Q)$ 的导数为 0 的点），于是得

$$\sum_{r=0}^{Q} P(r) = \frac{k}{k+h} \tag{6.15}$$

由于 k, h 及 $P(r)$ 都已知，所以从中可解出 Q。

现在我们利用公式（6.15）重新求解例 1。由已知

$$k = 700, h = 400, \frac{k}{k+h} \approx 0.637$$

又因为

$$P(0) = 0.05, P(1) = 0.10$$

$$P(2) = 0.25, P(3) = 0.35$$

$$\sum_{r=0}^{2} P(r) = 0.4 < 0.637 < \sum_{r=0}^{3} P(r) = 0.75$$

这说明订货量 Q 应在 2 与 3 之间，但 Q 更靠近于 3，因此订货量应为 3 千本。结论与前述解法相同。

例 2 某商店拟出售某种商品，其成本为每单位 50 元，售价 70 元；如不能售出，必须减价为 40 元，这时一定可以售出。已知售货量 r 的概率服从泊松分布

$$P(r)=\frac{e^{-\lambda}\lambda^{r}}{r!}.$$

其中 λ 为平均售出数。根据过去经验，平均售出数为 6 个单位，即 $\lambda=6$。求该商店的订货量应为多少单位。

解 该商店的缺货损失，每单位为 70-50 = 20；滞销损失，每单位为 50-40 = 10，即 $k=20,h=10$，故

$$\frac{k}{k+h}=\frac{20}{20+10}\approx 0.667$$

今记 $F(Q)=\sum_{r=0}^{Q}P(r)$，而 $P(r)=\dfrac{e^{-6}\cdot 6^{r}}{r!}$，查泊松分布表得

$$F(6)=\sum_{r=0}^{6}\frac{e^{-6}\cdot 6^{r}}{r!}=0.6063$$

$$F(7)=\sum_{r=0}^{7}\frac{e^{-6}\cdot 6^{r}}{r!}=0.7440$$

因为

$$F(6)<\frac{k}{k+h}<F(7)$$

而 $F(6)$ 的数值更接近于 0.667，所以订货量应为 6 个单位。

6.4.2 需求是随机离散变量的存储策略模型二

现在考虑这样的问题，假设在一个时间内：

（1）原有存储量为 I（在这一时间内为常数）；

（2）订购费为 C_3，订购物单价为 k，订货量为 Q；

（3）单位存储费为 C_1，缺货费为 C_2；

（4）需求量为 r，其概率 $P(r)$ 已知。

这一时期开始时，订货量为 Q，存储量达到 $I+Q$，这时所需的各种费用为：

订货费：

$$C_3+kQ$$

存储费：

当 $r\le I+Q$ 时，未能售出的部分需付存储费；当 $r>I+Q$ 时，不需付存储费，所以存储费的期望值为

$$\sum_{r\le I+Q}C_1\left(I+Q-r\right)P\left(r\right)$$

缺货费：

当 $r > I+Q$ 时，$(r-I-Q)$ 部分需付缺货费，所以缺货费的期望值为

$$\sum_{r>I+Q} C_2(r-I-Q)P(r)$$

记 $S = I+Q$（它表示存储达到的水平），则上述 3 项费用的期望值之和为

$$C(S) = C_3 + k(S-I) + \sum_{r\le S} C_1 P(r)(S-r) + \sum_{r<S} C_2 P(r)(r-S) \qquad (6.16)$$

于是问题归结为：S 取何值时，$C(S)$ 最小，即

$$\min C(S)$$

我们仿照模型一的办法解此模型。为此，

（1）将需求量 r 的随机取值按大小顺序排列为

$$r_0 < r_1 < \cdots < r_i < r_{i+1} < \cdots < r_m$$

$$\Delta r_i = r_{i+1} - r_i (i = 0,1,\cdots,m-1)$$

（2）S 只从 r_i 中取值。当 S 取值为 r_i 时，记为 S_i，

$$\Delta S_i = S_{i+1} - S_i = r_{i+1} - r_i = \Delta r_i > 0$$

$$\begin{aligned}
\Delta C(S_i) &= C(S_{i+1}) - C(S_i) \\
&= C_3 + k(S_{i+1} - I) + \sum_{r\le S_{i+1}} C_1 P(r)(S_{i+1} - r) + \sum_{r>S_{i+1}} C_2 P(r)(r - S_{i+1}) \\
&\quad - C_3 - k(S_i - I) - \sum_{r\le S_i} C_1(r)(S_i - r) - \sum_{r>S_i} C_2 P(r)(r - S_i) \\
&= k\Delta S_i + C_1 \Delta S_i \sum_{r\le S_i} P(r) - C_2 \Delta S_i \sum_{r>S_i} P(r) \\
&= k\Delta S_i + C_1 \Delta S_i \sum_{r\le S_i} P(r) - C_2 \Delta S_i [1 - \sum_{r\le S_i} P(r)] \\
&= k\Delta S_i + (C_1 + C_2)\Delta S_i \sum_{r\le S_i} P(r) - C_2 \Delta S_i
\end{aligned}$$

令 $\Delta C(S_i) = 0$，并注意到 $\Delta S_i \ne 0$，则得

$$\sum_{r\le S} P(r) = \frac{C_2 - k}{C_1 + C_2} \qquad (6.17)$$

记 $N = \dfrac{C_2 - k}{C_1 + C_2}$，并称其为临界值。由于需求是离散的，即 r 不连续取值，S 也不连续取值，所以（6.17）不一定成立。为此选取使不等式

$$\sum_{r\le S_i} P(r) \ge N$$

成立的 S_i 中的最小值作为 S，则订货量为

$$Q = S - I$$

例 3　某厂拟用塑料做原料制成产品出售，已知每箱塑料购价为 800 元，订购费 $C_3 = 60$ 元，存储费每箱 $C_1 = 40$ 元，缺货费每箱 $C_2 = 1015$ 元，原有存储量 $I = 10$ 箱。又已知对原料需求的概率为

$$P(30)=0.20,\quad P(40)=0.20$$
$$P(50)=0.40,\quad P(60)=0.20$$

试求该厂订购原料的最佳量。

解 先计算临界值

$$N=\frac{1015-800}{1015+40}\approx 0.204$$

再选取使不等式$\sum_{r\le S_i}P(r)\ge N$成立的最小S_i作为S。

$$P(30)=0.20<0.204$$
$$P(30)+P(40)=0.20+0.20=0.40>0.204$$

所以$S_i=40$作为S，最后得订货量为

$$Q=S-I=40-10=30\text{（箱）}$$

例 4 假定在国际市场上每年对我国某种出口商品的需求量是随机变量ξ（单位：吨），它服从［2000，4000］的均匀分布。设售出这种商品 1 吨，可为国家挣得外汇 3 万元，但假如销售不出而积压于仓库，则每吨需保养费 1 万元。应组织多少货源，才能使国家的收益最大?

解 若以y为组织某年出口的此种商品量（显然可以只考虑$2000\le y\le 4000$的情况），则收益（单位：万元）为

$$\eta=H(\xi)=\begin{cases}3y & \xi\ge y\\ 3\xi-(y-\xi) & \xi<y\end{cases}$$

因为ξ的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2000} & x\in[2000,4000]\\ 0 & x\notin[2000,4000]\end{cases}$$

于是收益的期望值为

$$\begin{aligned}E\eta&=\int_{-\infty}^{+\infty}H(x)f(x)\mathrm{d}x=\frac{1}{2000}\int_{2000}^{4000}H(x)\mathrm{d}x\\&=\frac{1}{2000}\int_{2000}^{y}(4x-y)\mathrm{d}x+\frac{1}{2000}\int_{y}^{4000}3y\mathrm{d}x\\&=\frac{1}{1000}(-y^2+7000y-4000000)\end{aligned}$$

此式当$y=3500$时达到最大，因此，组织 3500 吨此种商品为最好的策略。

习题六

1.有一批糖果，现从中随机地取 16 袋，称得质量(单位:克)如下:

506 508 499 503 504 510 497 512 514 505 493 496 506 502 509 496

设袋装糖果的质量近似地服从正态分布，试求总体均值和标准差的置信度为 0.95 的置信区间(提示:利用 Matlab 命令调用格式:[mu, sigma, muci, sigmaci]=normfit(X, alpha))。

2.考虑一个在时刻 $t=0$ 有 K 个不同姓氏的 N 个人构成的封闭社会系统，假设所有后代都随父姓，为回答以下问题，请作出必要的假设并建立一个数学模型。

（1）x 代以后姓氏的分布规律；

（2）某一姓氏消失的规律；

（3）一个姓氏存在的平均时间。

3.假定早晨上班期间人们随机地进入一座摩天大楼里的电梯，结果有 n 部电梯在每一层楼都停下来，各下了一两位乘客。

（1）讨论改善这种状况的计划。

（2）怎样衡量改善的程度?

（3）怎样才能模拟这种状况以便决定采取什么计划?

4.估计一个有限封闭区域内的生物数量，例如湖里的鱼或森林里的松鼠的数量，一种方法是抓到一些个体并给它们做上标记或系上标签，再放回原处，设想抓到了 x 只动物，做好标记后释放回去，过一段时间后，抓到了 n 只这类动物并发现其中有 y 只有标记。怎样估计这一地区该类动物的总数量 N 呢?估计出的数量的准确度如何?建立一个模拟模型来检查你的答案。对于 x 和 n 值的选取你有何建议?

5.三分球远投是整个篮球战术的一个组成部分，也是篮球比赛中的一个重要得分手段，但是如果盲目远投会给对方抢篮板球后反攻乃至得分造成机会，试从得失分的角度探讨三分球战术的原则，研究如下情况：现球权在自己一方，而对方领先不超过 2 分，除本次进攻外估计所剩时间不会超过两个回合（完成一次进攻算一个回合），作为指导，你建议本次进攻队员投 3 分球还是 2 分球?

第7章 差分方程建模

差分方程反映的是关于离散变量的取值与变化规律。通过建立一个或几个离散变量取值所满足的平衡关系，从而建立差分方程。具体而言，就是针对要解决的目标，引入系统或过程中的离散变量，根据实际背景的规律、性质、平衡关系，建立离散变量所满足的平衡关系等式，从而建立差分方程。通过求出和分析方程的解，或者分析得到方程解的特殊性质（平衡性、稳定性、渐近性、振动性、周期性等），从而把握这个离散变量的变化过程的规律，进一步再结合其他分析，得到原问题的解。

差分方程模型有着广泛的应用。实际上，连续变量可以用离散变量来近似和逼近，从而微分方程模型就可以近似于某个差分方程模型。差分方程模型有着非常广泛的实际背景。在经济、金融、保险领域、生物种群的数量结构规律分析、疾病和病虫害的控制与防治、遗传规律的研究等许多方面都有着非常重要的作用。可以这样讲，只要牵涉到关于变量的规律、性质，就可以适当地用差分方程模型来表现与分析求解。

在实际建立差分方程模型时，往往要将变化过程进行划分，划分成若干时段，根据要解决问题的目标，对每个时段引入相应的变量或向量，然后通过适当假设，根据事物系统的实际变化规律和数量相互关系，建立每两个相邻时段或几个相邻时段或者相隔某几个时段的量之间的变化规律和运算关系（即用相应设定的变量进行四则运算或基本初等函数运算或取最运算等）等式（可以多个并且应当充分、全面地反映所有可能的关系），从而建立起差分方程。或者对事物系统进行划分，划分成若干子系统，在每个子系统中引入恰当的变量或向量，然后分析建立起子过程间的这种量的关系等式，从而建立起差分方程。在这里，过程时段或子系统的划分方式是非常重要的，应当结合已有的信息和分析条件，从多种可选方式中挑选易于分析、针对性强的划分，同时，对划分后的时段或子过程，引入哪些变量或向量都是至关重要的，要仔细分析、选择，尽量扩大对过程或系统的数量感知范围，包括对

已有的、已知的若干量进行结合运算、取最运算等处理方式，目的是建立起简洁、深刻、易于求解分析的差分方程。在后面我们所举的实际例子中，这方面的内容应当重点体会。

差分方程模型是一种重要的数学模型。对它的应用也应当遵从一般的数学建模的理论与方法原则。同时注意与其他数学模型方法结合起来使用，因为一方面建立差分方程模型所用的数量、等式关系的建立都需要其他的数学分析方式来进行；另一方面，由差分方程获得的结果有可以进一步进行优化分析、满意度分析、分类分析、相关分析等等。

7.1　差分方程的基本知识

7.1.1　差分算子

设数列 $\{x_n\}$，定义差分算子 $\Delta:\Delta x_n = x_{n+1} - x_n$ 为 x_n 在 n 处的向前差分。$\Delta x_n = x_n - x_{n-1}$ 为 x_n 在 n 处的向后差分。

以后若无特别说明均指向前差分。

进一步定义 Δx_n 的差分：$\Delta(\Delta x_n) = \Delta^2 x_n$，

称之为在 n 处的二阶差分，它反映的是增量的增量。

类似地，可定义在n处的k阶差分为：$\Delta^k x_n = \Delta\left(\Delta^{k-1}(x_n)\right)$。

7.1.2　不变算子、平移算子

记 $Ex_n = x_{n+1}$，$Ix_n = x_n$，称 E 为平移算子，I 为不变算子，则有：

$$\Delta x_n = Ex_n - Ix_n = (E - I)x_n$$

所以 $\Delta = E - I$。于是

$$\Delta^k x_n = (E-I)^k x_n = \sum_{i=0}^{k}(-1)^{k-i}\mathrm{C}_k^i E^i x_n = \sum_{i=0}^{k}(-1)^{k-i}\mathrm{C}_k^i x_{n+i} \tag{7.1}$$

这表明 x_n 在 n 处的 k 阶差分由 x_n 在 $n, n+1, \cdots, n+k$ 处的取值线性决定。

反之，由 $\Delta x_n = x_{n+1} - x_n$ 得：　$x_{n+1} = \Delta x_n + x_n$；$\Delta^2 x_n = x_{n+2} - 2x_{n+1} + x_n$，从而得：

$$x_{n+2} = 2x_{n+1} - x_n + \Delta^2 x_n。$$

这个关系表明：第n+2项可以用前两项以及相邻三项增量的增量来表示和计算。即一个数列的任意一项都可以用其前面的k项和包括这项在内的 $k+1$ 项增量的增量的增量……第 k 层增量所构成。

由 $\Delta^k x_n = \sum_{i=0}^{k-1}(-1)^{k-i}\mathrm{C}_k^i x_{n+i} + x_{n+k}$ 得

$$x_{n+k} = \Delta^k x_n - \sum_{i=0}^{k-1}(-1)^{k-i} C_k^i x_{n+i} \tag{7.2}$$

可以看出：x_{n+k} 可以由 $x_n, \Delta x_n, \cdots, \Delta^k x_n$ 的线性组合表示出来。

7.1.3 差分方程

由 x_n 以及它的差分所构成的方程

$$\Delta^k x_n = f\left(x_n, \Delta x_n, \cdots, \Delta^{k-1} x_n\right) \tag{7.3}$$

称为 k 阶差分方程.

由（7.1）式可知（7.3）式可化为

$$x_{n+k} = F\left(x_n, x_{n+1}, \cdots, x_{n+k-1}\right) \tag{7.4}$$

故（7.4）也称为 k 阶差分方程（反映的是未知数列 $\{x_n\}$ 任意一项与其前面 k 项之间的关系）。

由（7.1）和（7.2）可知，（7.3）和（7.4）是等价的。

我们经常用的差分方程的形式是（7.4）式。

7.1.4 差分方程的解与有关概念

（1）如果 x_n 使 k 阶差分方程（7.4）总成立，则称 x_n 为方程（7.4）的解。

（2）如果 $x_n = \bar{x}$（$\bar{x}$ 为常数）是（7.4）的解，则称之为（7.4）的平衡解或平衡点。平衡解可能不止一个。平衡解的基本意义是：设 x_n 是（7.4）的解，考虑 x_n 的变化性态，其中之一是极限状况，如果 $\bar{x} = \lim\limits_{n\to\infty} x_n$，则方程（7.4）两边取极限（$x$ 就在这里面）。

（3）如果（7.4）的解 x_n 使得 $x_n - \bar{x}$ 既不是最终正的，也不是最终负的，则称 x_n 为关于平衡点 $\bar{x}$ 的振动解。

（4）如果令 $y_n = x_n - \bar{x}$，则方程（7.4）会变成

$$y_{n+k} = G\left(y_n, y_{n+1}, \cdots, y_{n+k-1}\right) \tag{7.5}$$

则 $y = 0$ 称为（7.5）的平衡点。

（5）如果（5）的所有解是关于 $y = 0$ 振动的，则称 k 阶差分方程（7.5）是振动方程。如果(7.5）的所有解是关于 $y = 0$ 非振动的，则称 k 阶差分方程（7.5）是非振动方程。

（6）如果（5）有解 y_n，使得对任意大的 N_y 有

$$\mathop{Sup}_{n\ge N_y}\left|y_n\right| > 0 \tag{7.6}$$

则称 y_n 为正则解（即不会从某项后全为零）。

（7）如果方程（4）的解 x_n 使得

$$\bar{x} = \lim_{n\to\infty} x_n \tag{7.7}$$

则称 x_n 为稳定解。

7.1.5　差分算子的若干性质

（1）$\Delta(\alpha x_n+\beta y_n)=\alpha\Delta(x_n)+\beta\Delta(y_n)$

（2）$\Delta\left(\dfrac{x_n}{y_n}\right)=\dfrac{1}{y_{n+1}y_n}(y_n\Delta x_n-x_n\Delta y_n)$

（3）$\Delta(x_n y_n)=y_{n+1}\Delta x_n+x_n\Delta y_n$

（4）$\sum\limits_{k=a}^{b}y_{k+1}\Delta x_k=x_{b+1}y_{b+1}-x_a y_a+\sum\limits_{k=a}^{b}x_k\Delta y_k$

（5）$x_n=E^n x_0=(\Delta+I)^n x_0=\sum\limits_{i=0}^{n}C_n^i\Delta^i x_0$

7.2　差分方程常用解法与性质分析

7.2.1　常系数线性差分方程的解

方程 $a_0x_{n+k}+a_1x_{n+k-1}+\cdots+a_kx_n=b(n)$　（7.8）

其中$a_0,a_1,\cdots,a_k$为常数，称方程（7.8）为常系数线性方程。

又称方程 $a_0x_{n+k}+a_1x_{n+k-1}+\cdots+a_kx_n=0$　（7.9）

为方程（7.8）对应的齐次方程。

如果（7.9）有形如$x_n=\lambda^n$的解，代入方程中可得

$$a_0\lambda^k+a_1\lambda^{k-1}+\cdots+a_{k-1}\lambda+a_k=0 \tag{7.10}$$

称方程（7.10）为方程（7.8）、（7.9）的特征方程。

显然，如果能求出（7.10）的根，则可以得到（7.9）的解。

结论：

（1）若（7.10）有k个不同的实根，则（7.9）有通解

$$x_n=c_1\lambda_1^n+c_2\lambda_2^n+\cdots+c_k\lambda_k^n$$

（2）若（7.10）有m重根λ，则通解中有构成项：

$$(\overline{c}_1+\overline{c}_2n+\cdots+\overline{c}_mn^{m-1})\lambda^n$$

（3）若（7.10）有一对单复根$\lambda=\alpha\pm\mathrm{i}\beta$，令

$$\lambda=\rho\mathrm{e}^{\pm i\varphi},\quad \rho=\sqrt{\alpha^2+\beta^2},\quad \varphi=\arctan\frac{\beta}{\alpha},$$

则（7.9）的通解中有构成项 $\overline{c}_1\rho^n\cos\varphi n+\overline{c}_2\rho^n\sin\varphi n$。

（4）若（7.10）有m重复根 $\lambda=\alpha\pm\mathrm{i}\beta$，$\lambda=\rho\mathrm{e}^{\pm i\varphi}$，则（7.9）的通项中有构成项

$$(\overline{c}_1+\overline{c}_2n+\cdots+\overline{c}_mn^{m-1})_1\rho^n\cos\varphi n+(\overline{c}_{m+1}+\overline{c}_{m+2}n+\cdots+\overline{c}_{2m}n^{m-1})\rho^n\sin\varphi n$$

综上所述，由于方程（7.10）恰有k个根，从而构成方程（7.9）的通解中必有k个独立的任意常数。通解可记为$\overline{x}_n$。

如果能得到方程（7.8）的一个特解x_n^*，则（7.8）必有通解

$$x_n = \overline{x}_n + x_n^* \tag{7.11}$$

（7.8）的特解可通过待定系数法来确定。

例如：如果$b(n) = b^n P_m(n)$，$P_m(n)$为n的多项式，则当b不是特征根时，可设成形如$b^n q_m(n)$形式的特解，其中$q_m(n)$为m次多项式；如果b是r重根时，设特解$b^n n^r q_m(n)$，将其代入（7.8）中确定出系数即可。

7.2.2 二阶线性差分方程组

设$Z(n) = \begin{pmatrix} x_n \\ y_n \end{pmatrix}, A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，形成向量方程组

$$Z(n+1) = AZ(n) \tag{7.12}$$

则

$$Z(n+1) = A^n Z(1) \tag{7.13}$$

（7.13）即为（7.12）的解。

为了具体求出（7.13），需要求出A^n，这可以用高等代数的方法计算。常用的方法有：

（1）如果A为正规矩阵，则A必可相似于对角矩阵，对角线上的元素就是A的特征值，相似变换矩阵由A的特征向量构成：

$$A = p^{-1}Ap \text{ , } A^n = p^{-1}A^n p$$

所以 $Z(n+1) = p^{-1}A^n pZ(1)$。

（2）将A分解成$A = \xi\eta'$，ξ, η为列向量，则有

$$A^n = (\xi\eta')^n = \xi\eta' \cdot \xi\eta' \cdots \xi\eta' = (\xi\eta')^{n-1} A$$

从而，$Z(n+1) = A^n Z(1) = (\xi' \eta)^{n-1} AZ(1)$。

（3）或者将A相似于约旦标准形的形式，通过讨论A的特征值的性态，找出A^n的内在构造规律，进而分析解$Z(n)$的变化规律，获得它的基本性质。

7.2.3 关于差分方程稳定性的几个结果

（1）k阶常系数线性差分方程（7.8）的解稳定的充分必要条件是它对应的特征方程（7.10）所有的特征根$\lambda_i, i = 1, 2, \cdots, k$满足$|\lambda_i| < 1$。

（2）一阶非线性差分方程

$$x_{n+1} = f(x_n) \tag{7.14}$$

的平衡点$\overline{x}$由方程$\overline{x} = f(\overline{x})$决定，将$f(x_n)$在点$\overline{x}$处展开为泰勒形式

$$f(x_n) \approx f'(\overline{x})(x_n - \overline{x}) + f(\overline{x}) \qquad (7.15))$$

故有$|f'(\overline{x})| < 1$时，（7.14）的解$\overline{x}$是稳定的；$|f'(\overline{x})| > 1$时，方程（7.14）的平衡点$\overline{x}$是不稳定的。

7.3 差分方程建模举例

差分方程建模方法的思想与一般数学建模的思想是一致的，也需要经历背景分析、确定目标、预想结果、引入必要的数值表示（变量、常量、函数、积分、导数、差分、取最等）概念和记号、几何形式（事物形状、过程轨迹、坐标系统等），也就是说要把事物的性态、结构、过程、成分等用数学概念、原理、方法来表现、分析、求解。当然，由于差分方程的特殊性，首先应当把系统或过程进行特别分解，形成表现整个系统的各个部分的离散取值形式，或形成变化运动过程的时间或距离的分化而得到离散变量。然后通过内在的机理分析，找出变量所能满足的平衡关系、增量或减量关系及规律，从而得到差分方程。另外，有时有可能通过多个离散变量的关系得到我们关心的变量的关系，这实际上建立的是离散向量方程，它有着非常重要的意义。有时还需要找出决定变量的初始条件。有时还需要将问题适当分成几个子部分，分别求解。

7.3.1 种群生态学中的虫口模型

在种群生态学中，考虑像蚕、蝉这种类型的昆虫数目的变化，它的变化规律是：每年夏季这种昆虫成虫产卵后全部死亡，第二年春天每个虫卵孵化成一个虫子。建立数学模型来表现虫子数目的变化规律。

应当考虑更具体的产卵分布状况，以 P_n 表示第n年的成虫数目，c 表示平均每个虫产卵的个数，则有 $P_{n+1}=cP_n$，这是一种简单模型。

如果进一步分析，由于成虫之间会有争斗以及传染病、天敌等的威胁，第 $n+1$ 年的成虫数会减少，如果考虑减少的主要原因是虫子之间的两两争斗，由于虫子配对数为

$$\frac{1}{2}P_n\left(P_n-1\right)\approx\frac{1}{2}P_n^2$$

故减少数应当与它成正比，从而有

$$P_{n+1}=cP_n-bP_n^2$$

这个模型可化成 $x_{n+1}=\lambda x_n\left(1-x_n\right)$，这是一阶非线性差分方程。这个模型的解的稳定性可以用相应一阶差分方程的判断方法，即（7.14）式来获得。

如果还考虑其他影响成虫孵卵及成活的因素的定量关系，这个模型在此基础上仍可进一步改进，更加符合实际情形。这种关系一方面可以通过机理分析，确定减少量与影响因素的定量关系，另一方面也可以用统计的方法来线性估计影响程度。或

者还可以用影响曲线的方法来直观表现影响的比例关系、周期关系、增量关系等等。

7.3.2 具有周期性运动过程的差分方程模型

建立差分方程描述振动台上的乒乓球垂直运动的方程，即把运动过程中的某些离散变化取值的变量的变化规律表现出来。

假设乒乓球与振动台之间的振动恢复系数为α，$\alpha \le 1$。振动台台面的上下位移是$-\beta\sin\bar{\omega}t$，乒乓球初始时刻在离台面垂直距离为H处为自由落体运动，$\beta << H$。又假设t_j为第j次碰撞时刻，第j次碰撞前的速度为$-u(t_j)$，碰撞后的速度为$v(t_j)$。假设

$$u(t_{j+1}) = v(t_j)$$

振动台台面的运动速度为

$$\omega(t) = \frac{\mathrm{d}}{\mathrm{d}t}(-\beta\sin\bar{\omega}t) = -\beta\cos\bar{\omega}t$$

又记$\phi = \tilde{\omega}t$，$v_j = \dfrac{2\tilde{\omega}v(t_j)}{g}$，则有

$$t_{j+1} - t_j = \frac{2v(t_j)}{g},\ \tilde{\omega}(t_{j+1} - t_j) = \frac{2\tilde{\omega}v_j}{g}, \phi_{j-1} - \phi_j = v_j \tag{7.16}$$

另外，由碰撞规律分析可知

$$v_{j+1} - \omega(t_{j+1}) = \alpha\left[-u_{j+1} + \omega(t_{j+1})\right]$$

该式经简化处理后可得：

$$v_{j+1} = \alpha v_j - \gamma\cos(\phi_j + v_j) \tag{7.17}$$

由（7.16）和（7.17）式联立可得二阶差分非线性方程组

$$\begin{cases} \phi_{j+1} - \phi_j = v_j \\ v_{j+1} = \alpha v_j - \gamma\cos(\phi_j + v_j) \end{cases}$$

7.3.3 人口控制与预测模型

（1）问题的提出

人口数量的发展变化规律及特性可以用偏微分方程的理论形式来表现和模拟。但在实际应用中不是很方便，需要建立离散化的模型，以便于分析、应用。人口数量的变化取决于诸多因素，比如：女性生育率、死亡率、性别比、人口基数等。试建立离散数学模型来表现人口数量的变化规律。

（2）模型假设

假设一　以年为时间单位记录人口数量，年龄取周岁。

假设二　某地区最大年龄为m岁。

假设三　第t年为i岁的人数为$x_i(t), i = 1, 2, \cdots, m$；$t = 1, 2, \cdots$。

注：这个数量指标是整个问题分析、表现的目标和载体，我们的目的就是找出这些变量的变化规律、内在的普遍联系。

假设四　t 年为 i 岁的人口平均死亡率为 $d_i(t)$，即这一年中 i 岁人口中死亡数与基数之比 $d_i(t)=\dfrac{x_i(t)-x_{i+1}(t+1)}{x_i(t)}$，即

$$x_{i+1}(t+1)=\left[1-d_i(t)\right]x_i(t), i=1,2,\cdots,m-1;\quad t=1,2,\cdots$$

假设五　第 t 年 i 岁女性的生育率：即每位女性平均生育婴儿数为 $b_i(t)$，$[i_1,i_2]$ 为生育区间。$k_i(t)$ 为第 t 年 i 岁人口的女性比（占全部 i 岁人口数）。

由此可知，第 t 年出生的人数为

$$f(t)=\sum_{i=i_1}^{i_2}b_i(t)k_i(t)x_i(t)$$

假设六　第 t 年婴儿的死亡率为 $d_{00}(t)$，则

$$x_0(t)=\left[1-d_{00}(t)\right]f(t)$$

假设七　$h_i(t)=\dfrac{b_i(t)}{\sum_{i=i_1}^{i_2}b_i(t)}=\dfrac{b_i(t)}{\beta(t)}$ 表示 i 岁女性总生育率，则 $b_i(t)=\beta(t)h_i(t)$。如果假设t年后女性出生率保持不变，则

$\beta(t)=b_{i1}(t)+b_{i1+1}(t)+\cdots+b_{i2}(t)=b_{i1}(t)+b_{i1+1}(t+1)+\cdots+b_{i2}(t+i_2-i_1)$

可见，$\beta(t)$ 表示每位妇女一生中平均生育的婴儿数，称之为总和生育率，它反映了人口变化的基本因素。

（3）模型建立

根据上面的假设

$$\begin{aligned}
x_1(t+1)&=\left[1-d_0(t)\right]x_0(t)\\
&=\left[1-d_0(t)\right]\left[1-d_{00}(t)\right]f(t)\\
&=\left[1-d_0(t)\right]\left[1-d_{00}(t)\right]\sum_{i=i_1}^{i_2}b_i(t)k_i(t)x_i(t)\\
&=\left[1-d_0(t)\right]\left[1-d_{00}(t)\right]\beta(t)\sum_{i=i_1}^{i_2}h_i(t)k_i(t)x_i(t)\\
&=\beta(t)\sum_{i=i_1}^{i_2}b_i'(t)x_i(t)
\end{aligned}$$

$$x_2(t+1)=\left[1-d_2(t)\right]x_1(t)$$

……

$$x_m(t+1)=\left[1-d_{m-1}(t)\right]x_{m-1}(t)$$

为了全面系统地反映一个时期内人口数量的状况，令

$$x(t)=\left[x_1(t) \quad x_2(t) \quad \cdots \quad x_m(t)\right]^{\mathrm{T}}$$

$$A(t)=\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1-d_1(t) & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1-d_2(t) & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & 1-d_{m-1}(t) & 0 \end{bmatrix}_{m\times n}$$

$$B(t)=\begin{bmatrix} 0 & 0 & b_{i_k}(t) & \cdots & b_{i_s}(t) & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}_{m\times n}$$

则此向量满足方程 $x(t+1)=A(t)x(t)+\beta(t)B(t)x(t)$，即

$$x(t+1)=\left[A(t)+\beta(t)B(t)\right]x(t) \tag{7.18}$$

这是一阶差分方程，其中$\beta(t)$是可控变量，$x(t)$是状态变量，并且关于$\beta(t)$和$x(t)$都是线性的，故称其为双线性方程。

（1）模型分析

在稳定的社会环境下，死亡率、生育模式、女性比例、婴儿存活率是可以假设为不变的，故$A(t)=A$，$B(t)=B$为常数矩阵。从而，

$$x(t+1)=\left[A+\beta(t)B\right]x(t) \tag{7.19}$$

只要总生育率$\beta(t)$确定下来，则人口的变化规律就可以确定下来。为了更全面地反映人口的有关信息，下面再引入一些重要的指标：

（1）人口总数：$N(t)=\sum\limits_{i=0}^{m}x_i(t)$

（2）人口平均年龄：$R(t)=\dfrac{1}{N(t)}\sum\limits_{i=0}^{m}i\cdot x_i(t)$

（3）平均寿命：$S(t)=\sum\limits_{j=0}^{m}j\exp\left[-\sum\limits_{i=0}^{j}d_i(t)\right]$，这里假定从第$t$年分析，如果以后每年的死亡率是不变的，即 $d_i(t)=d_{i+1}(t+1)=\cdots$，则$\sum\limits_{i=0}^{j}d_i(t)$表示$t$年出生的人活到第$j$+1年期间的死亡率，这也表明其寿命为$j$岁，$j=1,2,\cdots,m$。而$\exp\left[-\sum\limits_{i=0}^{j}d_i(t)\right]$表示寿命。

通过求出$x(t)$的变化规律，就可以对上面引入的3个指标进行更具体的分析，从而对人口的分布状况、变化趋势、总体特征等有科学的认识和把握。具体求解分析不再进行。

7.3.4 金融问题的差分方程模型

7.3.4.1 贷款还款模型

（1）问题的提出

设有一笔 p 万元的商业贷款，如果贷款期是 n 年，年利率是 r_1，今采用月还款的方式逐月偿还，建立数学模型计算每月的还款数是多少？

（2）模型分析

在整个还款过程中，每月还款数是固定的，而待还款数是变化的，找出这个变量的变化规律是解决问题的关键。

（3）模型假设

假设贷款后第 k 个月后的欠款数是 A_k 元；月还款为 m 元；月贷款利息为 $r=\dfrac{r_1}{12}$。

（4）模型建立

关于离散变量 A_k，考虑差分关系 $A_k+rA_k=A_{k+1}+m$，即

$$A_{k+1}=A_k(1+r)-m \tag{7.20}$$

这里已知 $A_0=100000$，$A_{24}=0$。

（5）模型求解

令 $B_k=A_k-A_{k-1}$，则 $B_k=B_{k-1}(1+r)=B_1(1+r)^{k-1}$，所以

$$\begin{aligned}A_k&=A_0+\sum_{i=1}^{k}B_i\\&=A_0+B_1\left[1+(1+r)+\cdots+(1+r)^{k-1}\right]\\&=A_0(1+r)^k-\frac{m}{r}\left[(1+r)^k-1\right],\quad k=0,1,2,\cdots\end{aligned}$$

这就是差分方程（7.20）的解。把已知数据 A_0，r 代入 $A_{12n}=0$ 中，可以求出月还款额 m。例如 $A_0=10000$，$r=0.0052125$，$n=2$ 时，可以求出：$m=444.356$ 元。

(6) 模型的拓广

拓广分析包括条件的改变、目标的改变、某些特殊结果等。如果令 $A_k=A$，则 $A=\dfrac{m}{r}$，并且当 $A_0=\dfrac{m}{r}$ 时，总有 $A_k=\dfrac{m}{r}$，即表明：每月只还上了利息。只有当 $A_0<\dfrac{m}{r}$ 时，欠款余额逐步减少，并最终还上贷款。

7.3.4.2 养老保险模型

（1）问题的提出

养老保险是保险中的一种重要险种，保险公司提供不同的保险方案供以选择，分析保险品种的实际投资价值。也就是说，如果已知所交保费和保险收入，按年或按月计算实际的利率是多少？也就是说，保险公司需要用你的保费实际获得至少多

少利润才能保证兑现你的保险收益?

例如:假设每月交费200元至60岁开始领取养老金,男子若25岁起投保,届时养老金每月为2282元;如35岁起保,届时月养老金为1056元;试求出保险公司为了兑现保险责任,每月至少应有多少投资收益率?这也就是投保人的实际收益率。

(2) 模型假设

这应当是一个过程分析模型问题。过程的结果在条件一定时是确定的。整个过程可以按月进行划分,因为交费是按月进行的。假设投保人到第k月止所交保费及收益的累计总额为F_k,每月收益率为r,用p表示60岁之前每月交费数,用q表示60岁之后每月领取数,N表示停交保险费的月份,M表示停领养老金的月份。

(1)模型建立

在整个过程中,离散变量F_k的变化规律满足:

$$\begin{cases} F_{k+1} = F_k(1+r) + p, \ k = 0, 1, \cdots, N-1 \\ F_{k+1} = F_k(1+r) - q, \ k = N, \cdots, M \end{cases}$$

这里F_k实际上表示从保险人开始交纳保险费以后,保险人账户上的资金数值,我们关心的是,在第M个月时,F_M能否为非负数?如果为正,则表明保险公司获得收益;如为负数,则表明保险公司出现亏损。当为零时,表明保险公司最后一无所有,表明所有的收益全归保险人,把它作为保险人的实际收益。从这个分析来看,引入变量F_k,很好地刻画了整个过程中资金的变化关系,特别是引入收益率r,虽然它不是我们所求的保险人的收益率,但是从问题系统环境中来看,必然要考虑引入另一对象:保险公司的经营效益,以此作为整个过程中各种量变化的表现基础。

(2)模型计算

以25岁起保为例。假设男性平均寿命为75岁,则有

$p = 20, q = 2282; N = 420, M = 600$,初始值为$F_0$,我们可以得到

$$F_k = F_0(1+r)^k + \frac{p}{r}\left[(1+r)^k - 1\right], k = 0, 1, 2, \cdots, N$$

$$F_k = F_N(1+r)^{k-N} - \frac{q}{r}\left[(1+r)^{k-N} - 1\right], k = N+1, \cdots, M$$

在上面两式中,分别取$k = N$和$k = M$并利用$F_M = 0$可以求出

$$(1+r)^M - \left(1 + \frac{q}{p}\right)(1+r)^{M-N} + \frac{q}{p} = 0$$

利用数学软件或利用牛顿法通过变成求出方程的根为:$r = 0.00485$。

用同样的方法可以求出:35岁和45岁起保所获得的月利率分别为:

$$r = 0.00461, \quad r = 0.00413$$

习题七

1.金融公司支付基金的流动模型：某金融机构设立一笔总额为 \$540 万的基金，分开放置在位于A城和B城的两个公司，基金在平时可以使用，但每周末结算时必须确保总额仍为 \$540 万。经过一段时间运行，每过一周，A城公司有10%的基金流动到B城公司，而B城公司则有12%的基金流动到A城公司。开始时，A城公司基金额为 \$260 万，B城公司基金额为 \$280 万。试建立差分方程模型分析：两公司的基金数额变化趋势如何？进一步要求，如果金融专家认为每个公司的支付基金不能少于 \$220 万，那么是否需要在什么时间将基金做专门调动来避免这种情况？

2.某保险公司推出与养老结合的人寿保险计划，其中介绍的例子为：如果40岁的男性投保人每年交保险费1540元，交费期20岁至60岁，则在他生存期间，45岁时（投保满5年）可获返还补贴4000元，50岁时（投保满10年）可获返还补贴5000元，其后每隔5年可获增幅为1000元的返还补贴。另外，在投保人去世或残废时，其受益人可获保险金20000元。试建立差分方程模型分析：若该投保人的寿命为76岁，其交保险费所获得的实际年利率是多少？寿命若为74岁时，实际年利率又是多少？

第 8 章　计算机仿真建模

8.1　计算机仿真建模概述

在一定假设条件下，利用数学运算模拟系统的运算，称为数学模拟。现代的数学模拟都是在计算机上进行的，因此称为计算机模拟，又称计算机仿真。模拟分为静态模拟和动态模拟。数值积分中的蒙特卡罗（Monte Carlo）方法就是典型的静态模拟。动态模拟分为连续系统模拟和离散系统模拟，连续系统模拟研究系统的状态随时间连续变化的情形，其模型一般是微分方程模型。建模时首先确定系统的连续状态变量，然后将它在时间上进行离散化处理，并由此模拟系统的运行状态。离散系统模拟讨论的是系统状态只在一些离散时间点上，由于随机事件驱动而发生变化，其模型一般用流程图或网络来表示。

为了模拟系统必须设置一个模拟时钟将时间从一个时刻向另一个时刻推进，并且可随时反映系统时间的当前值。模拟时间推进方式有两种：下次事件推进法和均匀间隔时间推进法。模拟离散系统常用下次事件推进法，其过程是：置模拟时间初值为0，跳到第一个事件发生的时刻，计算系统的状态，产生未来事件并加入到队列中去；跳到下一事件，计算系统的状态……重复这一过程直到满足某个终止条件为止。例如“海港系统的卸载货物问题”：假设某海港在任意时刻只允许一艘船卸载货物，由于船到达的时间是离散的，每艘船卸载货物的时间是随机的，因此，时间推进只能是以一艘船卸载货物从开始到完成，然后跳到下一艘船开始卸载货物。

一般来讲，在下列情况中，计算机模拟能有效地解决问题：

（1）难以用数学公式表示的系统，或者没有建立和求解数学模型的有效方法；

（2）虽然可以用解析的方法解决问题，但数学分析与计算过于复杂，此时计算机模拟可能提供简便可行的求解方法；

（3）希望能在较短的时间内观察到系统发展的全过程，以估计某些参数对系统行为的影响；

（4）难以在实际环境中进行实验和观察时，计算机模拟是唯一可行的方法，例如太空飞行的研究；

（5）需要对系统或过程进行长期运行比较，从大量方案中找到最优方案。

模拟连续系统常用均匀间隔时间推进法。

8.2　Monte Carlo 方法

8.2.1　Monte Carlo 方法的基本思想

Monte Carlo 方法源于世界著名的赌城——摩纳哥的 Monte Carlo。它是计算机仿真的基础，是基于对大量事件的统计结果来实现一些确定性问题的计算。此方法的基本思想是：为了计算某些量，先构造概率模型（如随机变量、随机过程等），利用概率模型的某些数字特征，恰好重合于所需计算的量，因此数字特征又可以通过实验，用统计方法求出它们的值，这些值作为要求的量的近似值。

如 1777 年法国科学家蒲丰（Buffon）提出计算 π 的一种方法：

在平面上，画出等距离 d（$d>0$）的一些平行线（图 8-1），向平面上随机投掷一长为 $l(l\le d)$ 的针。一共投掷 n 次（设 n 是一个相当大的数），共有 m 次与平行线相交，则针与平行线相交的概率是 $\frac{m}{n}$。另一方面，设 M 为针的中点，x 表示针投在平面上与 M 最近的直线距离，θ 表示针与此直线的夹角，则显然 $0\le x\le\frac{d}{2}$，$0\le\theta\le\pi$，当 $0\le x\le\frac{l}{2}\sin\theta$ 时，针与平行线相交。由几何概率知针与平行线相交的概率为 $p=\frac{2}{\pi d}\int_0^{\pi}\frac{l}{2}\sin\theta\mathrm{d}\theta=\frac{2l}{\pi d}$（图 8-2）。按照大数定理，当 n 充分大时，应有 $\frac{m}{n}\approx\frac{2l}{\pi d}$，由此知

$$\pi\approx\frac{2nl}{md}$$

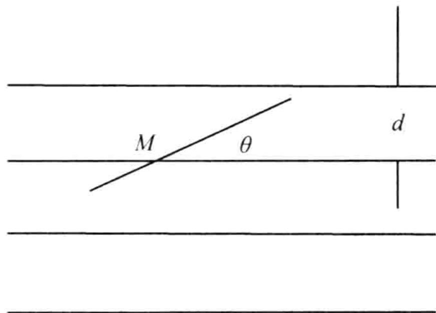


图 8-1

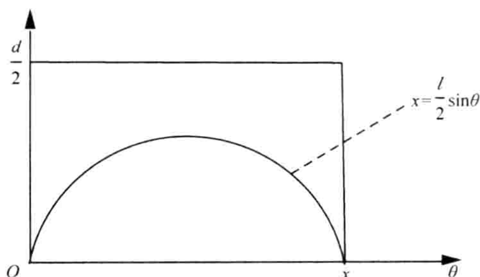


图 8-2

当然，这种方法求π需要在n充分大时才有效。因而用人工重复试验的方法基本上是行不通的，只有用计算机工具才使得有效使用上述方法成为可能，同时也给Monte Carlo方法带来了新的生机。特别是von Newmann和Uiam等人为研制原子弹，用计算机模拟链式反应过程，对该方法的发展起了巨大的推动作用。

8.2.2 Monte Carlo方法求积分的规则

任何一个积分都可看作某个随机变量的期望值，因此，可以用这个随机变量的平均值来近似它。

设要求的积分$I=\int_D f(p)\mathrm{d}p$，其中，D表示积分区域，取D上任意联合概率密度函数$\rho(p)$，令

$$F(p)=\begin{cases}\dfrac{f(p)}{\rho(p)}, & \rho(p)\neq 0\\ 0, & \rho(p)=0\end{cases}$$

则$I=\int_D F(p)\rho(p)\mathrm{d}p=E[F(p)]$，$p$的分布密度函数为$\rho(p)$。设$P_1, P_2, \cdots, P_n$为$D$上按$\rho(p)$选取的随机样本，只要积分$I$是有限值，则$I$是随机变量$F(p)$的数学期望值，其算术平均值$\bar{F}_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}F(p_i)$就是积分的近似。

选取$\rho(p)$最简单的方法是取区域D上的均匀分布，设D的体积（二维则为面积）为V_D，则

$$\rho(p)=\begin{cases}\dfrac{1}{V_D}, & p\in D\\ 0, & p\notin D\end{cases}$$

于是 $F(p)=V_D f(p)$，$I\approx\bar{F}_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}F(p_i)=\dfrac{V_D}{n}\sum\limits_{i=1}^{n}f(p_i)=V_D\bar{f}_n$，这就是求多重积分的Monte Carlo方法。利用中心极限定理易于证明当$n\to\infty$，抽样算术平均$\bar{F}_n$以概率1收敛到I，即$p\left\{\lim\limits_{n\to\infty}\bar{F}_n=I\right\}=1$，当$X$的概率分布为$\rho(x), a\le x\le b$时，则

$$E[f(X)]=\int_a^b f(x)\rho(x)dx。$$

特别地，当X为$(0,1)$上的均匀分布时，$\rho(x)=1$，此时

$$\int_0^1 f(x)\mathrm{d}x=E[f(X)]\approx\frac{1}{n}\sum_{k=1}^{n}f(x_k)$$

当计算在任一区间(a,b)上的积分时要先作代换$x=a+(b-a)u$化为$(0,1)$上的积分

$$\int_a^b f(x)\mathrm{d}x=(b-a)\int_0^1 f(a+(b-a)u)\mathrm{d}u\approx\frac{b-a}{n}\sum_{k=1}^{n}f(a+(b-a)u_k)$$

$u_k\ (k=1,2,\cdots,n)$ 为 $(0,1)$ 上的随机数。

8.2.3　随机数与随机变量的抽样

随机变量的抽样即是用适当的方法产生符合某种概率分布的随机变量的样本点。与之密切相关的是随机数。

最常用的随机数是在 $[0,1]$ 内均匀分布的随机数。通常可利用随机数表，将其存入计算机备用，也可以直接用随机数发生器产生一些 $(0,1)$ 内的二进制数。但考虑到随机数往往很多，且计算机的内存有限，实际中常用的是用计算机中的 Random 命令（或不同语言中的相应命令）生成一种称为“伪随机数”的数。虽然这还不是真正意义下的随机数，但由于这些方法均经过统计试验，从而具有较好的性质，实践中尽可放心使用。

利用随机数或伪随机数可以得到不同随机变量的抽样。

1.均匀分布的抽样

直接用 Random 命令（C 语言、Mathmation 语言中均适用）产生的随机数即为服从均匀分布的抽样。

2.离散分布的抽样

设 X 为离散型随机变量，$p_i=P(x=x_i)$，则当产生的随机数 r 满足

$$\sum_{i=1}^{k-1}p_i<r\le\sum_{i=1}^{k}p_i$$

时，取 $x=x_i$，这样产生的样本点即为服从相应分布的随机变量的样本点。事实上，由

$$P(x=x_i)=\frac{\sum_{i=1}^{k}p_i-\sum_{i=1}^{k-1}p_i}{1}=p_i$$

即知。

特别地，对 Poisson 分布 $P(x=n)=\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$，只要产生的随机数 r 满足

$$\sum_{k=0}^{n-1}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}<r<\sum_{k=0}^{n}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

取 $x=n$，即为服从 Poisson 分布的随机变量的样本点。

类似地，可得出二项分布、二点分布的抽样。

3.连续型分布的抽样

设连续型随机变量ξ的分布密度为$f(x)$，则分布函数$y=F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t$。若能从中解出$x=F^{-1}(y)$，即可产生出连续型随机变量$\xi$的抽样。具体做法是：首先产生随机数$r$（$\xi$服从均匀分布），则$\eta=F^{-1}(\xi)$即为分布函数$F(x)$的随机变量。

事实上，

$$P(\eta<x)=P\left(F^{-1}(\xi)<x\right)=P\left(\xi<F(x)\right)=F(x)\left(0\le F(x)\le 1\right)$$

即η为分布函数$F(x)$的随机变量。

例如，对指数分布，分布密度为$f(x)=\lambda\mathrm{e}^{-\lambda x}$，$x\ge 0$，$\lambda>0$，则

$$y=F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=-\int_{0}^{x}\mathrm{e}^{-\lambda t}d(-\lambda t)=-\mathrm{e}^{-\lambda t}\Big|_{0}^{x}=1-\mathrm{e}^{-\lambda t}$$

于是

$$F^{-1}(y)=-\frac{1}{\lambda}\ln(1-y),0\le y<1\text{或}F^{-1}(r)=-\frac{1}{\lambda}\ln r,0\le r<1\text{。}$$

上述抽样方法称为直接抽样，但上述方法并不总是可行的。例如，分布函数$y=F(x)$的反函数未必总是可以求出的，甚至$F(x)$也难以求出。于是对于一些特殊的分布函数可用近似方法得到抽样。

4.近似抽样——正态分布的抽样

对于正态分布，$F(x)$就难以求出，更无法求出F^{-1}，下面考虑近似抽样。

设$\xi_1,\xi_2,\cdots,\xi_n$是n个相互独立的在$[0,1]$上服从均匀分布的随机变量，$E(\xi_i)=\frac{1}{2}$，$D(\xi_i)=\frac{1}{12}$，由中心极限定理知随机变量$\eta_n=\left(\sum_{i=1}^{n}\xi_i-\frac{n}{2}\right)\Big/\sqrt{\frac{n}{12}}$，当$n\to\infty$时趋于标准正态分布。

于是，可在$(0,1)$内随机地取$r_1,r_2,\cdots,r_n$，当n充分大时，可用$x=\left(\sum_{i=1}^{n}r_i-\frac{n}{2}\right)\Big/\sqrt{\frac{n}{12}}$近似代替标准正态分布$N(0,1)$。

为保证精度，要取n充分大。一般情况下取$n\ge 10$即可，但由于式中分母有 12，故取 12 即得

$$\begin{aligned}x&=\left(\sum_{i=1}^{12}r_i-6\right)\Big/\sqrt{\frac{12}{12}}=r_1+r_2+\cdots+r_{12}-6\\&=r_1-(1-r_2)+r_3-(1-r_4)+\cdots+r_{11}-(1-r_{12})\end{aligned}$$

或

$$x=\sum_{i=1}^{6}\left(r_{2i}-r_{2i-1}\right)$$

即为服从正态分布 $N(0,1)$ 的样本点。

由概率论知识，若令 $\eta=\sigma\xi+\alpha$，则当 $\xi\sim N(0,1)$ 时，$\eta\sim N(\alpha,\sigma)$。由此知，当 r 为服从正态分布 $N(0,1)$ 的抽样时，$\sigma r+\alpha$ 为服从正态分布 $N(\alpha,\sigma)$ 的样本点。

几个常见分布的抽样：

（1）Maxwell 分布 $f(x)=\dfrac{2\beta^{\frac{3}{2}}}{\sqrt{\pi x}}\mathrm{e}^{-\beta x},x\ge 0,\beta>0$ 的抽样：

随机产生两个随机数 $r_1,r_2\,(r_2\ne 0)$，若 $r_1^2\le -\mathrm{e}r_2\ln r_2$，则取 $x=-\dfrac{3}{2\beta}\ln r_2$，否则，继续产生 $r_1,r_2\,(r_2\ne 0)$。

（2）二维正态分布 $f(x,y)=\dfrac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}$ 的抽样：

$$x=\sqrt{-2\ln r_1}\cos 2\pi r_2,\quad y=\sqrt{-2\ln r_1}\sin 2\pi r_2\ (r_1\ne 0)。$$

（3）β 分布 $f(x)=\dfrac{(n+1)!}{m!(n-m)!}x^m(1-x)^{n-m},(0\le x\le 1)$ 的抽样：

随机产生 $n+1$ 个随机数，取第 $m+1$ 个最小数即为服从 β 分布的样本点。

（4）Γ 分布 $f(x)=\dfrac{\alpha^n}{(n-1)!}x^{n-1}\mathrm{e}^{-\alpha x}$ 的抽样：

随机产生 n 个随机数 $r_1,r_2,\cdots,r_n$，$x=-\dfrac{1}{\alpha}\ln(r_1\cdots r_n)$，$(r_1\cdots r_n>0)$ 为服从 Γ 分布的样本点。

（5）χ^2 分布密度函数为 $f(x)=\dfrac{1}{2_{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}},(x>0)$ 的抽样：

当 $y_i\,(i=1,2,\cdots,n)$ 为 n 个服从正态分布 $N(0,1)$ 的独立抽样的样本点时，则 $x=\sum_{i=1}^{n}y_i$ 为服从 χ^2 分布的样本点。

8.3　基于遗传蚁群算法求解 TSP 优化问题的研究

8.3.1　引言

蚁群算法（ant colony algorithm）是由 Dorigo、Maniezzo 和 Colorni 于 1991 年提出的，这是最初的 ACO 算法。该算法早期应用于旅行商问题（Traveling Salesman Problem，TSP）、指派问题、调度问题等问题求解。特征选择问题是一个 NP 问题，选择方法分为两大类： Filter 和 Wrappe。其主要思想是搜索一个接近近似解的最优

解。Filter 类主要采用具体独立归纳学习算法，其主要特点是对整个路径搜索速度较快，但算法在应用过程中易出现早熟现象，使得准确率较低；Wrapper 类同样也具有归纳学习算法的特点并以其性能作为评价和选择特征的标准，其主要特点是算法准确率较高，但每一次对子集的评价都要进行重组计算，计算复杂度高，不适合大规模数据的搜索与计算。结合 Filter 和 Wrapper 是两种互补的模式，两者的结合为研究蚁群算法提供了较大的考虑价值[1]。文献[2]是将权值引入到对信息素进行更新的过程中，通过重复迭代提高对全局的搜索能力从而获取最短路径。文献[3]是利用由信息素转移概率与分布估计算法相结合产生新个体来提高收敛速度和全局搜索能力以达到抑制早熟等缺陷的目的。文献[4]提出一种遗传蚁群算法，其思想是利用遗传算法先对全局进行搜索求解，而后利用蚁群算法对求得的解进行优化以得到最优解的目的。

本文提出改进遗传蚁群算法，采用 Filter 类和 Wrapper 类的共同特点，以互信息作为启发函数，利用蚁群算法来动态搜索特征子集解，并结合分类器的分类效果作为评价准则，通过对分类精度指导特征结点的信息素更新，通过对遗传算法产生一个较优解的同时动态改变交叉算子和变异算子，求解出较优路径上的信息素强度，然后再按照蚁群算法完成一次遍历后，对路径上信息素做局部信息素迭代更新和全局更新，从而增强了全局优化能力，加快了系统收敛速度，提高了搜索效率，避免了迭代过程中早熟现象的出现，抑制了冗余码的产生，使得求解精度更为准确从而有效地抑制了过早收敛现象的发生，改进了遗传蚁群算法在全局收敛的局限性，提高了收敛速度，实验结果表明该算法具有较高的效率。

8.3.2 遗传蚁群算法原理与特性分析

1.蚁群算法原理

蚁群算法是模仿真实世界蚁群的行为提出的。真实的蚂蚁在没有视觉的情况下，能够找到从食物源到蚁巢的最短路径；同时，它们能适应环境的改变，当原先最短的路径上出现障碍物时，能够发现一条新的最短路径。这是因为蚂蚁个体之间是通过一种被称为信息素的物质进行信息传递的。蚂蚁在运动过程中，不但能够在它所经过的路径上留下该物质，而且能够感知这种物质的存在及其强度，并朝着该物质强度高的方向移动，为此指导自己的运动方向。因此，由大量蚂蚁组成的蚁群集体行为表现出一种信息正反馈现象。某一路径上走过的蚂蚁越多，则后来的蚂蚁选择该路径的概率就越大。蚂蚁个体之间就是通过这种信息的交流进行路径的最优选择，从而达到搜索食物的目的。

2.遗传算法原理

遗传算法是一种仿生型优化算法，是生物界中自然选择和自然遗传机制的随机化自适应搜索算法。通过作用于染色体上的基因寻找最好的染色体进行求解。在遗传算法中，利用交叉算子和变异算子对解空间进行搜索，交叉算子是通过组合父辈个体的特性，产生新的个体。它和选择算子相结合后，是加快遗传算法信息交换的主要方法，增强了遗传算法的全局搜索性。然后再利用适应度函数对每个个体进行数值评价，淘汰低适应度的个体，对这个新种群进行下一轮的进化。每一个个体都表示所求问题的一个潜在隐性解，在遗传操作中将一代一代地向最优解进化。

3.蚁群算法模型

从蚂蚁集体觅食行为得到启发，Dorigo 等人提出蚁群优化算法。算法利用大量蚂蚁在移动过程中感知遗留在路径上的信息素来指导自己的运动等特点，通过协同寻优实现最短路径选择等行为。蚁群算法的这种寻优能力也可用于特征选择。将蚁群模型应用于特征选择时，将每个特征作为一个图结点，蚂蚁不必遍历所有特征，特征之间也不再以距离为评价因素，蚁群系统的特征选择模型构造完全图 G：构造一个 n 阶完全图 $G=(V,E')$。其中 n 为特征数，V 为顶点集，E' 为边集。在图 G 中，蚂蚁随机选择一个顶点作为搜索的起始点，蚂蚁通过遍历完全图 G 寻找一个包含最少结点数量并达到最小的分类错误率的特征子集。假设给定初始特征集合 $F=\{f_1,\cdots,f_n\}$，n 是用于搜索特征空间的蚂蚁数量,对于蚂蚁 k 都包含一个列表 L 用于存放蚂蚁 k 已选择的特征，$\tau(f_i)$ 表示某时刻特征结点 f_i 上剩余的信息素，在实际应用问题中可解释为特征 f_i 对数据集分类效果的重要度，信息素与特征值成正比，$\tau(f_i)$ 越大，蚂蚁选择该特征的可能性就越大。

4.蚁群算法特性分析

蚁群算法的优点主要体现在：

第一，蚁群算法是一种结合了分布式计算、正反馈机制和贪婪式搜索的算法，具有很强的搜索较优解的能力。正反馈能够快速地发现较优解，分布式计算避免了早熟收敛，而贪婪式搜索有助于在搜索过程中早期找出可接受的解决方案，缩短了搜索时间。

第二，蚁群算法具有很强的鲁棒性，对蚁群算法的模型稍加修改，就可以应用于其他问题。

第三，具有易于与多种启发式算法结合以改善算法的性能。

第四，可以不通过个体之间直接通信而是通过信息素进行合作，这样的系统具

有更好的可扩充性。由于随着系统中个体增加而增加的系统通信开销在这里将非常小。

该算法也存在一些缺陷：

第一，该算法一般需要较长的搜索时间。蚁群中各个个体的运动是随机的，通过信息交换能够向着最优路径进化，但是当群体规模较大时，很难在短时间内从大量杂乱无章的路径中找出一条较好的路径。这是因为在进化的初级阶段，各个路径上的信息量相差不明显。通过信息正反馈，使得较好路径上的信息量逐渐增大，经过较长一段时间，才能使得较好路径上的信息量明显高于其他路径上的信息量，随着这一过程的进行，差别越来越明显，从而最终收敛于较好的路径。这一过程一般需要较长的时间。

第二，该算法容易出现停滞现象，即搜索进行到一定程度后，所有个体所发现的解完全一致，不能对解空间进行进一步搜索，不利于发现更好的解。

8.3.3 遗传蚁群算法建立

1.特征权值蚁群优化算法

在加快收敛和降低停滞之间取得平衡是目前蚁群算法研究的难点之一。既要使算法的搜索空间尽可能大，以寻找那些可能存在最优解的解空间；同时又要充分利用群体内当前具有的有效信息，使算法的侧重点放在那些可能具有较高适应值的个体所在空间，从而以较大概率收敛到全局最优解。利用动态加权方法，根据蚂蚁所求解质量不同，赋予不同权重。根据权重的不同，对蚂蚁所经过的特征信息素自适应更新。为此，要判断蚁群是聚向一个或者仅仅几个解还是散布在搜索空间当中。可以通过计算本次迭代所有特征子集解的平均分类错误率 E_{avg} 和本次迭代最小特征子集解的分类错误率 $E_{\min}$ 之间的差别来判断蚁群的聚合度。可以推断，如果整个蚁群是聚合的，那么这个差别比相对分散的蚁群要小。因此可以将两者之间的差别作为衡量蚁群是否收敛的一个尺度。

蚁群构造特征解的 E_{avg} 为：

$$E_{\text{avg}} = \frac{1}{n}\sum_{k=1}^{na} E_k \tag{8.1}$$

E 是本次迭代中蚂蚁构造的特征子集解的分类错误率，在更新信息素时，对 E 少于 E_{avg} 的蚂蚁赋予一个相对较大的权值，越少赋予的权值越大，对于 E 多于或等于 E_{avg} 的蚂蚁赋予一个相对小的固定权值。因此蚂蚁 k 在本次迭代中所选择的特征上释放的信息素为：

$$\Delta\tau(f_i) = \begin{cases} Q\lambda_k / E_{\text{avg}} & f_i\text{被选择} \\ 0 & \text{否则} \end{cases} \tag{8.2}$$

其中的权值 λ_k 由下式求得：

$$\lambda_k=\begin{cases}(E_{\text{avg}}-E_k)/(E_{\text{avg}}-E_{\min}) & E_k<E_{\text{avg}}\\ 0 & E_k\ge E_{\text{avg}}\end{cases} \tag{8.3}$$

可以看到 E_k 越接近 $E_{\min}$ 时，权重 λ 越接近 1。根据公式（8.2）计算每次迭代的蚂蚁构造特征 E_{avg}，再根据个体蚂蚁构造特征解的 E 与 E_{avg} 之间的差异来对信息素进行相应的改变。这样就可以增加解空间的多样性，提高蚁群的全局搜索能力，避免局部收敛过快，同时也不会降低蚁群的搜索速度。

在蚁群算法中，蚂蚁构造解时选择下一结点的原则和信息素更新规则对蚁群最终能否搜索到最优解起到决定性作用。

2.遗传算法的建立

长度为 L 的 n 个二进制串 s_i（i=1, 2, 3, …, n）组成了遗传算法的初解群，也称为初始群体。在每个串中，每个二进制位就是个体染色体的基因。对群体执行的操作有三种：

第一种为选择操作，这是从群体中选择出较适应环境的个体。这些选中的个体用于繁殖下一代。故有时也称这一操作为再生。适应值比例选择是最基本的选择方法，其中每个个体被选择的期望数量与其值和群体平均适应值的比例有关。给定规模为 n 的群体 $P=\{a_1,a_2,a_3,\cdots,a_n\}$，个体 $a_j\in P$ 的适应值函数为 $f(a_j)$，其选择概率为

$$p_s(a_j)=f(a_j)\Big/\sum_{i=1}^{n}f(a_i) \tag{8.4}$$

第二种为交叉操作，这是在选中用于繁殖下一代的个体中，对两个不同的个体的相同位置的基因进行交换，从而产生新的个体。交叉操作一般可分三种，分别是单点交叉操作、多点交叉操作和一致交叉操作。其中一致交叉操作是目前所研究的交叉操作中最为核心的交叉操作之一。一致交叉操作就是染色体位串上的每一位按相同的概率进行随机均匀交叉。

第三种为变异操作，这是在选中的个体中，对个体中的某些基因执行异向转化。在位串 s_i 中，如果某位基因为1，产生变异时就是把它变成0，反之亦反之。个体发生变异的原始概率为 p_m，通过个体变异概率计算得出群体中个体发生变异的概率为

$$p_m(a_j)=1-(1-p_m)^L \tag{8.5}$$

8.3.4　遗传蚁群算法改进

1.启发函数与构造解

蚂蚁在搜索过程中总是希望最先选取最能满足评价函数的结点。初始时，所有

结点的信息素都是一致的，更需要启发函数对蚂蚁进行引导。在基本 TSP 问题中，启发函数定义为两个城市之间距离的倒数，这在对特征结点实施搜索时是不适用的。由于 Filter 类特征选择方法中，特征和类别之间的互信息能较好地反映特征的重要性，为此我们将每个特征结点的启发函数定义为该特征与类别的互信息，如下所示：

$$\eta_{f_i} = I(f_i:C) \quad (8.6)$$

f_i 表示条件特征，C 表示类别特征。

初始时每个蚂蚁被随机地放置在一个特征结点上，从该起点出发依据状态转移规则选择下一个特征。蚂蚁构造完解后，对 $k \in [1, n]$ 的所有特征解分别进行训练分类，得到分类错误率，并据此评估其优劣。单个蚂蚁选择特征的依据如公式（8.7）、（8.8）所示。

$$S = \begin{cases} \arg\max\{\tau(u)^{\alpha} * \eta(u)^{\beta}\}, & if(q < q_0),\ u \in allowed_k \\ p(s), \text{ otherwise}, & s \in allowed_k \end{cases} \quad (8.7)$$

$$p(s) = \frac{[\tau(s)]^{\alpha} * [\eta(s)]^{\beta}}{\sum\limits_{u \in allowed_k} [\tau(u)]^{\alpha} * [\eta(u)]^{\beta}} \quad (8.8)$$

其中：α，β 表示信息素和启发信息的相对重要性；$allowed_k$ 指当前蚂蚁 k 搜索所有可选特征集合；s 表示蚂蚁 k 当前选择的特征；u 表示与特征 s 邻接的特征；$allowed = \{0,1,2,\cdots,n-1\} - tabu_k$ 表示蚂蚁 k 下一步允许选择的所有结点，列表 $tabu_k$ 记录了当前蚂蚁 k 所走过的结点，当所有 n 个结点都写入到禁忌表 $tabu_k$ 中时，蚂蚁 k 就完成一次循环，所经过的路径就是问题的一个解。q_0 是均匀分布在区间 $[0,1]$ 的一个随机变量，用于确定规则两部分的相对重要性，即蚂蚁选择能使信息素和启发信息的乘积最大的特征和蚂蚁选择具有最大转移概率的特征两部分。以此保证蚂蚁利用以往信息的同时也具有开拓新搜索空间的能力，防止过快收敛到局部最优。

2.全局更新机制与局部更新机制

蚂蚁从结点 i 转移到结点 j 后，对路径 (i, j) 上的信息素按公式（8.9）更新

$$\tau_{ij}(t+1) = (1-\xi)\tau_{ij}(t) + \xi\,\Delta\tau_{ij}(t) \quad (8.9)$$

其中 $\xi \in (0,1)$，$\Delta\tau_{ij} = \Delta\tau(0)$ 是信息量的初值。这样可以有效地避免陷入局部最优，减小了已选择过的路径再次被选择的概率，从而有效地避免蚂蚁收敛到同一路径，进而提高了算法的全局搜索能力。

蚂蚁的搜索主要集中在当前循环为止所找出的最短路径的领域区，全局更新规则是在所有的蚂蚁都完成它们的路径之后执行，更新信息素公式是

$$\tau_{ij}(t+1)=\left(1-\xi\tau_{ij}(t)\right)\tau_{ij}(t)+\xi\tau_{ij}(t)\Delta\tau_{ij}(t) \quad (8.10)$$

$$\Delta\tau_{ij}^{k}=\begin{cases}\dfrac{Q}{L_{\min}} & 边(i,j)\in L_{\min}\\ -\dfrac{Q}{L_{\max}} & 边(i,j)\in L_{\max}\\ 0 & \text{otherwise}\end{cases} \quad (8.11)$$

其中，ξ 是信息素挥发系数，如对 ξ 做一定的变化，可以防止该路径上的信息素浓度无限地增加，可减少陷入局部最优化的可能，从而加快了收敛速度。$L_{\min}$，$L_{\max}$ 是蚂蚁所走过的最短和最长路径。在算法的初期阶段，让蚂蚁走过的各路径都有机会获得信息量的更新，进而提高蚂蚁选择相应路径的概率。随着某些路径上信息素的不断积累，在循环次数达到一定数目时，缩小了最优路径与最差路径的边之间的信息量差距，从而使蚂蚁的搜索行为集中于最优路径的附近。

3.变异操作与交叉操作

在变异操作过程中，由路径 L_0 变异到另一条路径 L_1 上，所采用逆转策略。在第 $1\sim n$ 个访问的城市中随机地选取第 j_1 次和第 j_2 访问的城市，在路径 L_0 中第 j_1 次到第 j_2 次访问城市之间的子路径以反方向插入，其余不变，得到新路径 L_1。

设 $L_0=1\,8\,3\,2\,6\,9\,5\,7\,4$，$j_1=3$，$j_2=6$，则 $L_1=1\,8\,9\,6\,2\,3\,5\,7\,4$。在交叉过程中，设在子串 $child2=1\,8\,3\,2\,6\,9\,5\,7\,4$ 中随机选择一个交叉区域为 $3\,2\,6\,9$，将子串 $child2$ 中的交叉区域加到子串 $child1=4\,7\,5\,9\,6\,2\,3\,8\,1$ 中对应的位置，删除子串 $child1$ 中已在 $child2$ 的交叉区中出现过的城市后，得到新子串 $child3=4\,7\,5\,3\,2\,6\,9\,8\,1$。

4.算法步骤与描述

在本文算法中，根据数据集设定特征子集个数 m 来避免不必要的搜索。对于任意一只蚂蚁选取 m 个特征后就完成一次搜索过程，m 为搜索周期，每个周期能够搜索出包含 m 个特征的特征子集的近似最优，然后 m 值以一定的比率增加。

Step1：初始化参数：令时间 $t=0$，循环次数 $N_c=0$，设置 $\eta_{ik}=d_{ik}$；$\tau_{ij}=0$；$\Delta\tau_{ij}=0$，并将出发点置于禁忌表 $tabu_k$ 中，给定 α,β,ρ 的值，边 $(i,j)\in R$。

Step2：当 *while* 不满足条件算法终止条件时，利用 *for* 循环将 m 只蚂蚁放置在初始结点上。

Step3：蚂蚁个体根据状态转移公式（8.6）计算的概率选择结点 j 并前进，此时 $j\in(C,tabu_k)$。

Step4：修改禁忌表中的指针，将结点 j 写入禁忌表中，让指针指向结点 j。

Step5：查找是否有未遍历的结点，如果 $k<m$，执行 $k++$ 后跳转到 Step3。// 每跳转一次，都要修改禁忌表中的指针，让其指向下一个新的结点。

Step6：计算 m 蚂蚁的路径长度 L，记录当前的最优解 L_{min}，求出第一次循环中得到的最优路径 L_k。

Step7：当蚂蚁遍历过所有结点后，就要对路径上的信息素按公式（8.7）、（8.8）、（8.9）更新。

Step8：当 $L_{min} \le L_k$ 时，更新 τ_{ik} 和 $\Delta\tau_{ik}$，并替换最优路径表。如果循环次数大于或等于循环最大次数，循环结束并输出最优路径长度 L_k。否则清空禁忌表，并跳转到 Step2。

8.3.5 体系评估

现以 TSP 为仿真对象。仿真实验从通用 TSPLIB 中选取三个对称型 TSP 实例进行仿真实验。开发环境为 Microsoft Windows XP，Matlab6.5，AMD 2.4GHz，1G 内存。其参数为 $\alpha=2$，$\beta=3$，$\xi=0.1$，$\tau_0=0.1$，$Q=100$，令蚂蚁数目与城市数目相等，实验进行了 50 次，每次进行 100 次循环，求 20 次基本蚁群算法和改进算法结果的最优解、最差解和平均解，见表 8-1 和表 8-2。

表 8-1　基本蚁群算法实验结果

问题	基本蚁群算法		
	最优解	最差解	平均解
Olive30	425.703	431.071	427.611
Eil51	429.885	449.593	436.007
Eil76	552.107	568.389	562.508
CHN144	31339.1	32799.5	31785.6

表 8-2　本文算法实验结果

问题	本文算法		
	最优解	最差解	平均解
Olive30	422.325	427.003	425.068
Eil51	426.000	429.102	428.061
Eil76	544.895	549.927	546.623
CHN144	30839.3	31331.3	31158.7

通过表 8-1 和表 8-2 的实验结果对比可知，本文算法所得最差值都要小于基本蚁群算法的最优值。这说明本文算法具有较好的全局搜索最优解的能力，加快了系统

收敛速度，避免了迭代过程中早熟现象的出现，抑制了冗余码的产生，使得求解精度更为准确，从而实现了动态优化求解过程。表 8-1 和表 8-2 所列举的是在参数相同情况下三组对称 TSP 基本蚁群算法与本文算法的比较。

为了验证本文算法的实效性，用基本蚁群算法和文献[8]与本文算法对 CHN144 进行对比实验，进行 50 次实验对比，最优解对应的收敛曲线如图 8-3 所示：

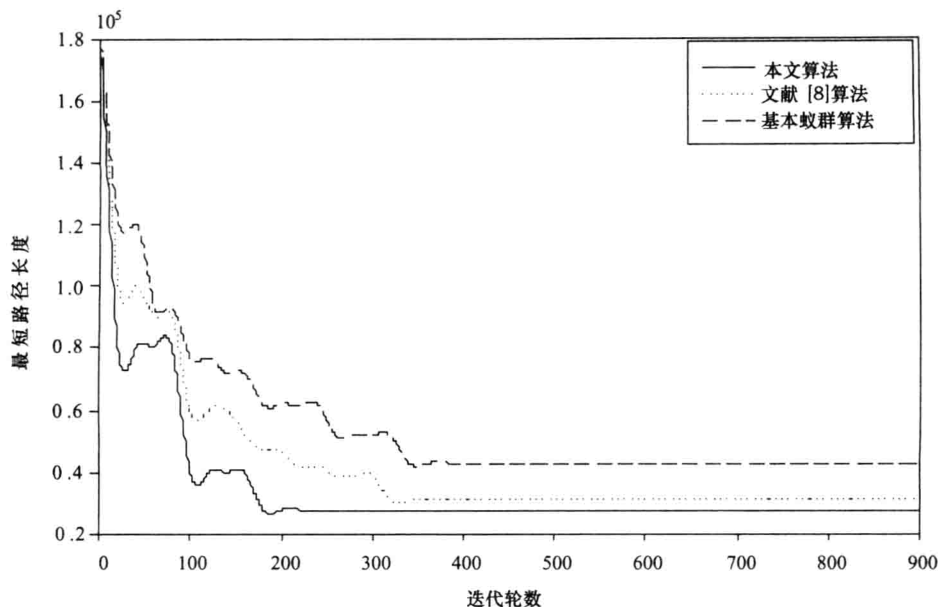


图 8-3 短路径收敛曲线对比图

图 8-3 描述了三种算法最短路径与迭代轮数之间的收敛曲线，从图 8-3 中可以看出初始时刻下，三种算法的收敛速度相当，由于本文算法是以行为特征为策略，在解决 TSP 问题时，解的多样性与解的质量明显优于其他两种算法。随着时间轮数的增加，本文算法收敛速度和优化性能高于文献[8]所给算法和基本蚁群算法。

8.3.6 结束语

本章将遗传蚁群算法求解最优化问题的优势应用于特征选择，并给出行为特征选择遗传蚁群算法优化过程，同时针对基本蚁群算法的局限性进行相应的改进，通过启发式函数、信息更新策略、交叉和选择策略对解的值进行优化，在一定程度上抑制了过早收敛现象，提高了对全局的搜索能力，有效地避免了陷入局部最优解，加快了收敛速度，最后通过 TSP 仿真实验验证了本书算法的可行性和有效性。

习题八

猎狗攻击问题

一只野兔在平面上沿着直线以速度v_1匀速前进，突然，一只猎狗攻击它，这只猎狗在侧后方以恒定速率v_2 $(v_2 > v_1)$追向该兔，狗的运动方向始终指向这只兔子。

（1）设开始时兔在A处，坐标为(x_0, y_0)，狗在原点O处，兔子沿平行于y轴直线奔跑，请建立数学模型求狗的运动轨迹。

（2）在如下数据下，计算并画出狗运动的轨迹，v_1=1 m/s, v_2=3 m/s, x_0=10 m, y_0=20 m。

第 9 章　几何建模

数学模型就是为了达到某种目的，用字母、数字及其他数学符号建立起来的等式或不等式以及图表、图像、框图等描述客观事物的特征及其内在联系的数学结构表达式。为了证明某两个区间（或区域）同胚，需要在它们之间寻找拓扑变换，但是一般情况下这种拓扑变换不易找到。本章将用几何的方法构建这种拓扑变换。

9.1　区间之间的同胚问题

设 I 为区间，2^I 表示 I 的所有子集作成的族，容易验证 2^I 为 I 的一个拓扑，因此 I 按 2^I 构成拓扑空间。

拓扑变换是指两个同胚的拓扑空间之间的连续的双射。两个拓扑空间同胚的充要条件是它们之间存在拓扑变换；区间是一种简单的拓扑空间；所以证明两个区间同胚归结为在它们之间找拓扑变换。

9.1.1　线段之间拓扑变换的构建

定理 9.1　若 $[a,b],[c,d]$ 都是闭区间，则它们同胚，即 $[a,b]\cong[c,d]$。

分析　欲证 $[a,b]\cong[c,d]$，只需找到它们之间的一个拓扑变换。

证明　（矩形对角线法）将 $[a,b],[c,d]$ 分别置于 x 轴和 y 轴，以它们为边构造矩形 $ABCD$（如图 9-1），对角线 AC 的方程是

$$y=\frac{d-c}{b-a}x+\frac{bc-ad}{b-a}$$

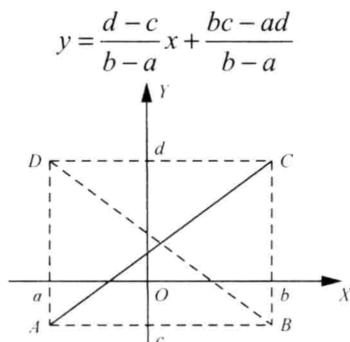


图 9-1

显然

$$T_1(x)=\frac{d-c}{b-a}x+\frac{bc-ad}{b-a}, x\in[a,b] \tag{9.1}$$

是双射且连续，所以是$[a,b]$到$[c,d]$的拓扑变换，因此 $[a,b]\cong[c,d]$。

注： 1. $T_1^{-1}(x)$是$[c,d]$到$[a,b]$的拓扑变换。

2.可以求出对角线BD的方程，由此确定$[a,b]$到$[c,d]$的另一个拓扑变换

$$T_2(x)=-\frac{d-c}{b-a}x+\frac{bd-ac}{b-a}, x\in[a,b]$$

3.在定理9.1中，把闭区间换为开区间，则$T_1(x)$和$T_2(x)$都是(a, b)到(c, d)的拓扑变换；$T_1^{-1}(x)$和$T_2^{-1}(x)$都是(c, d)到(a, b)的拓扑变换。

9.1.2 无端点线段与直线之间的拓扑变换的构建

定理 9.2 无穷区间$(-\infty, +\infty)$ 和开区间(a, b) 同胚，即$(-\infty, +\infty)\cong(a, b)$。

证法 1 （折线投影法）先将横轴上的开区间(a, b)向上映照到折线 $y=\left|x-\frac{a+b}{2}\right|$上，再从坐标为$A\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$的点将此折线（无两端）投射到横轴上（如图9-2）。

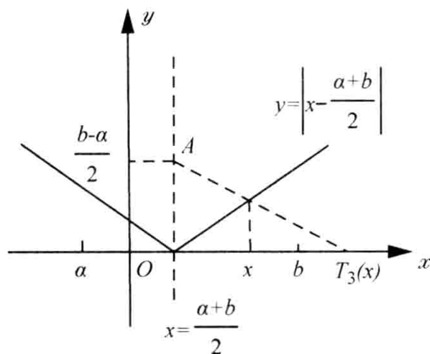


图 9-2

于是有

$$\frac{T_3(x)-x}{T_3(x)-\frac{a+b}{2}}=\frac{\left|x-\frac{a+b}{2}\right|}{\frac{b-a}{2}}$$

解得

$$T_3(x)=\frac{\frac{b-a}{2}x-\frac{a+b}{2}\left|x-\frac{a+b}{2}\right|}{\frac{b-a}{2}-\left|x-\frac{a+b}{2}\right|} \tag{9.2}$$

下证T_3是$(a,\ b)$到$(-\infty,\ +\infty)$的拓扑变换。

显然T_3是双射；由初等函数的连续性，$T_3(x)$在$\left(a,\ \dfrac{a+b}{2}\right)$和$\left(\dfrac{a+b}{2},\ b\right)$都连续，而

$$T_3\left(\frac{a+b}{2}-0\right)=\lim_{x\to\left(\frac{a+b}{2}\right)^-}\frac{\dfrac{b-a}{2}x+\dfrac{a+b}{2}\left(x-\dfrac{a+b}{2}\right)}{\dfrac{b-a}{2}+\left(x-\dfrac{a+b}{2}\right)}=\frac{a+b}{2}$$

$$T_3\left(\frac{a+b}{2}+0\right)=\lim_{x\to\left(\frac{a+b}{2}\right)^+}\frac{\dfrac{b-a}{2}x-\dfrac{a+b}{2}\left(x-\dfrac{a+b}{2}\right)}{\dfrac{b-a}{2}-\left(x-\dfrac{a+b}{2}\right)}=\frac{a+b}{2}$$

所以 $\lim\limits_{x\to\frac{a+b}{2}}T_3(x)=\dfrac{a+b}{2}=T_3\left(\dfrac{a+b}{2}\right)$，$T_3$在开区间$(a,\ b)$内连续。

由（9.2）式解得，

$$T_3^{-1}(x)=\begin{cases}\dfrac{4ax-(a+b)^2}{4(x-b)}\ , & x<\dfrac{a+b}{2}\\[2ex] \dfrac{4bx-(a+b)^2}{4(x-a)}\ , & \dfrac{a+b}{2}\le x\end{cases}$$

同法可证 $T_3^{-1}:(-\infty,+\infty)\to(a,\ b)$ 是连续的双射。所以T_3是$(a,\ b)$到$(-\infty,+\infty)$的一个拓扑变换。

证法 2 （圆弧投影法）将证法 1 中的折线改为以点 $C\left(\dfrac{a+b}{2},\ \dfrac{b-a}{2}\right)$ 为圆心，$\dfrac{b-a}{2}$为半径的下半开圆，保持投影方法不变（图 9-3）。

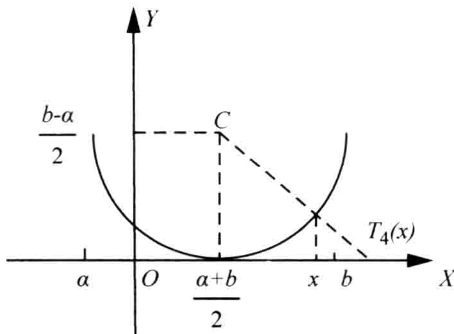


图 9-3

可以求得函数

$$T_4(x)=\frac{x-\frac{a+b}{2}}{\sqrt{1-\frac{4\left(x-\frac{a+b}{2}\right)^2}{(b-a)^2}}}+\frac{a+b}{2},x\in(a,\quad b) \tag{9.3}$$

解得

$$T_4^{-1}(x)=\frac{x-\frac{a+b}{2}}{\sqrt{1+\frac{4\left(x-\frac{a+b}{2}\right)^2}{(b-a)^2}}}+\frac{a+b}{2},x\in(-\infty,\quad +\infty) \tag{9.4}$$

易证T_4，T_4^{-1}都是连续的双射。因此T_4是(a, b)到$(-\infty, +\infty)$的拓扑变换。

注 在（9.3）式中分别以a、$b-a$代换$\frac{a+b}{2}$、$\frac{b-a}{2}$，立得$[a, b)$到$[a, +\infty)$的拓扑变换

$$T_5(x)=\frac{x-a}{\sqrt{1-\frac{(x-a)^2}{(b-a)^2}}}+a,x\in[a,\quad b)$$

又以b、$b-a$分别代换$\frac{a+b}{2}$、$\frac{b-a}{2}$，得到$(a, b]$到$(-\infty, b]$的拓扑变换

$$T_6(x)=\frac{x-b}{\sqrt{1-\frac{(x-b)^2}{(b-a)^2}}}+b,x\in(a,\quad b]$$

证法 3 （缠绕投影法）在证法 2 中，假如在作T_4时不是把开区间(a, b)正投影上去，而是均匀拉长$\frac{\pi}{2}$倍重叠到半圆上再作中心投影，则对$\forall x\in(a,\ b)$有

$$\frac{T_4(x)-\frac{a+b}{2}}{\frac{b-a}{2}}=\tan\frac{\pi}{b-a}\left(x-\frac{a+b}{2}\right)$$

解得

$$T_4(x)=\frac{b-a}{2}\tan\frac{\pi}{b-a}\left(x-\frac{a+b}{2}\right)+\frac{a+b}{2},x\in(a,\quad b) \tag{9.5}$$

解得

$$T_4^{-1}(x)=\frac{b-a}{\pi}\arctan\left(\frac{2x}{b-a}+\frac{a+b}{a-b}\right)+\frac{a+b}{2},x\in(-\infty,\quad +\infty)$$

显然，T_4, T_4^{-1} 都是连续双射，因此 T_4 是 (a, b) 到 $(-\infty, +\infty)$ 的拓扑变换。

证法 4　先证 $(a, b) \cong (-1, 1)$，再证 $(-\infty, +\infty) \cong (-1, 1)$，由同胚的传递性和对称性，得 $(-\infty, +\infty) \cong (a, b)$。

事实上，作变换 $T(x) = -\dfrac{a-b}{2}x + \dfrac{a+b}{2}$，显然，$T(x)$ 是 $(-1, 1)$ 到 (a, b) 的连续双射，$T^{-1}(x)$ 是 (a, b) 到 $(-1, 1)$ 的连续双射，所以 $T:(-1,1) \to (a,b)$ 是拓扑变换，因此 $(a,b) \cong (-1,1)$。

由证法 1，$f(x) = \dfrac{x}{1-|x|}$ 是 $(-1,1)$ 到 $(-\infty, +\infty)$ 的拓扑变换，因此，$(-\infty, +\infty) \cong (-1,1)$，由同胚的传递性和对称性，$(-\infty, +\infty) \cong (a, b)$。

显然，两个同胚的区间之间的拓扑变换是不唯一的，而且找法不同得到的拓扑变换可能不同。

9.1.3　无端点射线和无端点线段间拓扑变换的构建

定理 9.3　无穷区间 $(m, +\infty)$ 与开区间 (a, b) 同胚，即 $(m, +\infty) \cong (a, b)$。

证明　令 $\phi(x) = \mathrm{e}^x + m$，显然 $\phi:(-\infty, +\infty) \to (m, +\infty)$ 是拓扑变换。由证法 2，$T_4:(a, b) \to (-\infty, +\infty)$ 是拓扑变换，于是

$$\phi \cdot T_4:(a, b) \to (m, +\infty),$$

$$T_5(x) = (\phi \cdot T_4)(x) = \phi[T_4(x)] = \mathrm{e}^{\left[\frac{b-a}{2}\tan\frac{\pi}{b-a}\left(x-\frac{a+b}{2}\right)+\frac{a+b}{2}\right]} + m$$

是拓扑变换。

定理 9.3 的证法称为拓扑变换复合法。与其类似的定理还有：

定理 9.4　无穷区间 $(-\infty, n)$ 与开区间 (a, b) 同胚，即 $(-\infty, n) \cong (a, b)$。

以上几种拓扑变换有着重要的应用，我们以第一个为例来说明。

1. 在定积分数值计算中的应用

例　利用三点 Gauss-Legendre 求积公式

$$\int_{-1}^{1} f(x)\mathrm{d}x \approx \frac{5}{9}f\left(-\frac{\sqrt{15}}{5}\right) + \frac{8}{9}f(0) + \frac{5}{9}f\left(\frac{\sqrt{15}}{5}\right)$$

计算积分 $\displaystyle\int_0^1 \frac{\ln(1+x)}{1+x^2}\mathrm{d}x$。

解　设 $g(x) = \dfrac{\ln(1+x)}{1+x^2}, x \in [0, 1]$，根据拓扑变换（9.1），令

$$x = \frac{t}{2} + \frac{1}{2}, t \in [-1, 1],$$

则

$$g\left[\frac{t}{2}+\frac{1}{2}\right]=\frac{4\ln\left(\frac{t+3}{2}\right)}{t^2+2t+5},t\in[-1,\ 1],\mathrm{d}x=\frac{1}{2}\mathrm{d}t$$

于是

$$\int_0^1\frac{\ln(1+x)}{1+x^2}\mathrm{d}x=\int_0^1 g(x)\mathrm{d}x=\int_{-1}^1 f(t)\mathrm{d}t\approx\frac{5}{9}f\left(-\frac{\sqrt{15}}{5}\right)+\frac{8}{9}f(0)+\frac{5}{9}f\left(\frac{\sqrt{15}}{5}\right)$$

其中

$$f(t)=\frac{2\ln\frac{t+3}{2}}{t^2+2t+5}$$

2.推广二次函数的一个性质

性质 1 设二次函数 $f(x)=ax^2+bx+c(a\neq 0)$，若 $|f(x)|$ 在 $[-1,1]$ 上的最大值为 M，则 $M\geq\frac{1}{2}|a|$。

下面将这一性质推广到更一般的闭区间上去。

性质 2 设二次函数 $f(x)=ax^2+bx+c(a\neq 0)$，若 $|f(x)|$ 在闭区间 $[d,\ e]$ 上的最大值为 M，则 $M\geq\frac{(e-d)^2}{8}|a|$。

证明 由（9.1）式作拓扑变换 $x=\frac{e-d}{2}(t-1)+e$，$t\in[-1,1]$，则有

$$\begin{aligned}f\left[\frac{e-d}{2}(t-1)+e\right]&=a\left[\frac{e-d}{2}(t-1)+e\right]^2+b\left[\frac{e-d}{2}(t-1)+e\right]+c\\&=\frac{a(e-d)^2}{4}t^2+\frac{(e-d)(ad+ae+b)}{2}t+\frac{a(e+d)^2}{4}+\frac{b(e+d)}{2}+c\end{aligned}$$

由性质 1，$M\geq\frac{1}{2}\left|\frac{a(e-d)^2}{4}\right|=\frac{(e-d)^2}{8}|a|$

9.2 度量空间与其开球同胚问题

受定理 9.2 的启发，本节我们研究度量空间与其开球同胚问题。

定理 9.5 设 $\varepsilon>0,(a,\ b)\in R^2$，定义

$$A=\left\{(x_1,\ x_2)\,\middle|\,(x_1-a)^2+(x_2-b)^2<\varepsilon^2,(x_1,\ x_2)\in R^2\right\},$$

则 $A\cong R^2$。

证明 记 $B=\left\{(x_1,\ x_2)\,\middle|\,x_1^2+x_2^2<1,(x_1,\ x_2)\in R^2\right\}$，作映射 l^p $(p\geq 1)$ 使得

$$\varphi(x_1,\ x_2)=(\varphi(x_1),\ \varphi(x_2)),\quad(x_1,\ x_2)\in A$$

其中 $\varphi(x_1)=\dfrac{x_1-a}{\varepsilon}$, $\varphi(x_2)=\dfrac{x_2-b}{\varepsilon}$。

由于$\varphi(x_1),\varphi(x_2)$都是连续双射，所以$\varphi$（$x_1, x_2$）是连续双射。同理，$\varphi^{-1}$（$x_1, x_2$）也是连续双射，所以$\varphi$是拓扑变换。因此，$B\cong A$。

构造映射

$$\psi: R^2 \to B,\quad \psi(x_1, x_2)=(\psi(x_1),\psi(x_2)),\quad (x_1,\ x_2)\in R^2$$

其中$\psi(x_1)=\dfrac{x_1}{1+\sqrt{x_1^2+x_2^2}}$, $\psi(x_2)=\dfrac{x_2}{1+\sqrt{x_1^2+x_2^2}}$。

因为$\psi(x_1),\psi(x_2)$都是双射且连续，所以$\psi(x_1, x_2)$是连续的双射；同理，$\psi^{-1}(x_1, x_2)$也是连续的双射。可见ψ是拓扑变换，因此 $B\cong R^2$。由同胚的对称性和传递性，$A\cong R^2$。

一般地，在任意有限维度量空间 R^n 中有：

定理 9.6　设$\varepsilon>0, x_0\in R^n, B(x_0, \varepsilon)=\{x|d(x, x_0)<\varepsilon, x\in R^n\}$是 n 维度量空间 R^n 的任一开球，则 $R^n\cong B(x_0, \varepsilon)$。

证明　构造映射

$$f: B(x_0, \varepsilon)\to B(0, 1).\quad \forall x\in B(x_0, \varepsilon),\quad f(x)=\frac{1}{\varepsilon}(x-x_0)$$

其中 x=（$x_1, x_2 \cdots, x_n$），x_0=（ $x_0^{(1)}, x_0^{(2)}\cdots, x_0^{(3)}$ ），$x_0^{(i)}, x_i\in R, i$=1, 2, $\cdots, n$。

即

$$f(x_1, x_2\cdots, x_n)=\frac{x^1-x_0^{(1)}}{\varepsilon}, \frac{x^2-x_0^{(2)}}{\varepsilon},\ \cdots,\ \frac{x_n-x_0^{(n)}}{\varepsilon}$$

显然，f是单射

又$\forall z=(z_1, z_2 \ldots, z_n)\in B(0, 1)$，有 $d(z, 0)=\sqrt{\sum_{i=1}^{n} z_i^2}<1$。令 u=（$u_1, u_2 \cdots, u_n$），$u_i=\varepsilon z_i+x_0^{(1)}$, i=1, 2, $\cdots, n$，于是

$$\sqrt{\sum_{i=1}^{n}(u^i-x_0^{(i)})^2}=\varepsilon\sqrt{\sum_{i=1}^{n} z_i^2}<\varepsilon$$

故 u=（$u_1, u_2, \cdots, u_n$）$\in B$（u_0, ε）。因f（$u_1, u_2, \cdots, u_n$）$=z$，可见f又是满射，又$f(x_i)=\dfrac{x_i-x_0^{(i)}}{\varepsilon}$, i=1, 2, $\cdots, n$ 是连续函数，因此f是连续的双射。

同法可证f^{-1}（x）$=\varepsilon x+x_0$, $\forall x\in B$（0, 1）亦是连续的双射，所以f是拓扑变换，因此，B（0, 1）$\cong B$（x_0, ε）

构造映射 $g: R_n \to B(0,1)$，$g(x)=\dfrac{x}{1+\|x\|}$，$\|x\|=d(0,x)$，$\forall x\in R_n$

设 $x=(x_1, x_2, \cdots, x_n)$，$x_i\in R, i=1, 2, \cdots, n$，则

$$g(x_1, x_2, \cdots, x_n)=(g(x_1), g(x_2), \cdots, g(x_n))$$

其中 $\dfrac{x_i}{1+\sqrt{\sum\limits_{k=1}^{n} x_k^2}}$，$i$=1, 2, $\cdots$, n。

显然 g 是单射；对 $\forall z\in B(0,1)$，$z=(z_1, z_2, \cdots, z_n)$，$z_i\in R$，$i$=1, 2, $\cdots$, n。因 $d(z, 0)<1$，即 $\|z\|<1$，令 $u=\dfrac{z}{1-\|z\|}$，显然 $u\in R^n$，又 $g(u)=\dfrac{u}{1+\|u\|}=z$，所以 g 是满射。又因为 $g(x_i)$ 是连续函数，所以 $g(x)=\dfrac{x}{1+\|x\|}$ 是连续的双射。

同法可证 $g^{-1}(x)=\dfrac{x}{1+\|x\|}$，$\forall x\in B(0,1)$ 是连续的双射。可见 g 是拓扑变换，因此 $B(0,1)\cong R^n$。

由同胚的对称性和传递性，$R^n\cong B(x_0, \varepsilon)$。

9.3 赋范线性空间与其开球对等和同胚问题

本节将给出"n 维欧氏空间与它的开球对等"的一种证法；然后定义赋范线性空间的开球；在此基础上推出"赋范线性空间与它的任意一个开球对等"这一结论；进而得到赋范线性空间与它的开球同胚的充要条件。

9.3.1 预备知识

定义 9.1 设 A 和 B 是两个非空集合，若它们之间存在双射，则称这两个集合等势或对等，记作 $A\sim$B，对等的两个集合之间的双射称为对等映射。

定义 9.2 设 X 和 Y 是两个拓扑空间，若存在双射 $f:X\to Y$，并且 f 和它的逆映射 f^{-1} 都连续，则称 f 是 X 到 Y 的拓扑变换或同胚映射，并称 X 和 Y 同胚，记作 $A\cong B$。

定义 9.3 设 $\varepsilon>0$，$x_0\in R^n$，称以 x_0 为心、ε 为半径的邻域

$$B(x_0,\varepsilon)=\left\{x \mid d(x,x_0)<\varepsilon, x\in R^n\right\}$$

是 R^n 的开球。

定义 9.4 设 $x=(\xi_1,\xi_2,\cdots)$ 是实（或复）数列，如果 $\sum\limits_{k=1}^{\infty}|\xi_k|^p<\infty,(p>1)$，则称数列 $(\xi_1,\xi_2,\cdots)$ 是 p 次收敛数列，p 次收敛数列全体记为 l^p。

定义 9.5　$f(x)$是$[a,\ b]$上实值可测函数，$1 \le p < +\infty$，若$|f(x)|^p$是$[a,\ b]$上的 Lebesgue 可积函数，则称$f(x)$是$[a,\ b]$上p方可积函数，$[a,\ b]$上p方可积函数全体记为$L^p[a,\ b]$。

定理 9.7　$l^p(p>1)$按通常的线性运算及范数

$$\|x\|_p = \left(\sum_{k=1}^{\infty}|\xi_k|^p\right)^{\frac{1}{p}},\left(x=(\xi_1,\xi_2,\cdots)\in l^p\right)$$

称为赋范线性空间。

定理 9.8　当$p \ge 1$时，$L^p[a,\ b]$ 按通常的函数线性运算及范数

$$\|f\|_p = \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}},\ \left(f \in L^p[a,\ b]\right)$$

称为赋范线性空间。

定理 9.9　若$\phi_i(t)(i=1,2,\cdots,n)$是向量函数$\phi(t)$ 的分量，即

$$\phi(t)=\left(\phi_1(t),\phi_2(t),\cdots,\phi_n(t)\right)$$

则$\phi(t)$ 当且仅当$\phi_i(t)(i=1,2,\cdots,n)$都连续时连续。

9.3.2　R^n~B（x_0, ε）的一种证法

定理 9.10　n维欧氏空间R^n与它的任意一个开球对等，即$R^n \sim B(x_0,\varepsilon)$。

证明　构造映射

$$f:B(x_0\,,\varepsilon)\to B(0\,,1),\quad \forall\, x\in B(x_0\,,\varepsilon),\quad f(x)=\frac{1}{\varepsilon}(x-x_0)$$

其中　$x=(x_1\,,x_2\,\ldots,x_n)$，$x_0=\left(x_0^{(1)}\,,x_0^{(2)}\,\ldots,x_0^{(n)}\right),x_i\,,x_0^{(i)}\in R\,,i=1\,,2\,,\cdots,n$

即

$$f(x_1\,,x_2\,\ldots,x_n)=\left(\frac{x_1-x_0^{(1)}}{\varepsilon},\frac{x_2-x_0^{(2)}}{\varepsilon},\cdots,\frac{x_n-x_0^{(n)}}{\varepsilon}\right)$$

显然f是单射。

$\forall\, z=(z_1\,,z_2\,\ldots,z_n)\in B(0,\ 1)$，有$d(z,0)=\sqrt{\sum_{i=1}^{n}z_i^2}<1$令$u=(u_1\,,u_2\,\ldots,u_n)$，$u_i=\varepsilon z_i+x_0^{(i)}\,,i=1\,,2\,,\cdots,n$，于是

$$\sqrt{\sum_{i=1}^{n}\left(u_i-x_0^{(i)}\right)^2}=\varepsilon\sqrt{\sum_{i=1}^{n}z_i^2}<\varepsilon$$

故$u=(u_1\,,u_2\,\ldots,u_n)\in B(x_0\,,\ \varepsilon)$，又

$$f(u)=f(u_1,u_2\dots,u_n)=\left(\frac{u_1-x_0^{(1)}}{\varepsilon},\frac{u_2-x_0^{(2)}}{\varepsilon},\cdots,\frac{u_n-x_0^{(n)}}{\varepsilon}\right)$$

$$=\left(\frac{\varepsilon z_1}{\varepsilon},\frac{\varepsilon z_2}{\varepsilon},\cdots,\frac{\varepsilon z_n}{\varepsilon}\right)=(z_1,z_2,\cdots,z_n)=z$$

因此 f 是满射，所以 $B(0,\ 1)\sim B(x_0,\varepsilon)$。

构造映射

$$g:R^n\to B(0\ ,\ 1),\quad g(x)=\frac{x}{1+d(0,\ x)},\ \forall\, x\in R^n$$

设 $x=(x_1\ ,x_2\ ,\dots,x_n)$，$x_i\in R\ ,i=1\ ,2\ ,\cdots,n$ ，则

$$g(x_1\ ,x_2\ ,\dots,x_n)=(h(x_1),h(x_2),\cdots,h(x_n))$$

其中 $h(x_i)=\dfrac{x_i}{1+\sqrt{\sum\limits_{k=1}^{n}x_k^2}}$，$i=1,2,\cdots,n$。

显然 g 是单射；对 $\forall\, z\in B(0\ ,\ 1)\ ,z=(z_1\ ,z_2\ ,\cdots,z_n)$，$z_i\in R\ ,i=1\ ,2\ ,\cdots,n$ ，因 $d(z,0)=\sqrt{\sum\limits_{i=1}^{n}z_i^2}<1$ ，令 $u=\dfrac{z}{1-d(0,\ z)}$，显然 $u\in R^n$，又

$$g(u)=\frac{u}{1+d(0,\ u)}=\frac{z}{1-d(0,z)}\Bigg/\left(1+d\left(0,\ \frac{z}{1-d(0,z)}\right)\right)=z$$

所以 g 又是满射，故 $B(0,\ 1)\sim R^n$。

由对等的对称性和传递性，$R^n\sim B(x_0,\varepsilon)$。

9.3.3 赋范线性空间的开球

定义 9.6 设 X 是赋范线性空间，$\varepsilon>0$，$x_0\in X$，称以 x_0 为心、ε 为半径的邻域

$$B(x_0,\varepsilon)=\{x\,|\,\|x-x_0\|<\varepsilon,x\in X\}$$

是 X 的开球。

按此定义，定理 9.7 中的赋范线性空间 $l^p\ (p\geq 1)$ 的开球是

$$B(x_0,\varepsilon)=\left\{x\,|\,\|x-x_0\|_p<\varepsilon,x\in l^p\right\}$$

定理 9.8 中的赋范线性空间 $L^p[a,\ b]\ (p\geq 1)$ 的开球是

$$B(f_0,\varepsilon)=\left\{f\,|\,\|f-f_0\|_p<\varepsilon,f\in L^p[a,\ b]\right\}$$

9.3.4 主要结论

为了将定理 9.10 推广到一般的赋范线性空间中去，先探究下面的两个命题。

命题 9.1 定理 9.7 中的赋范线性空间 $l^p\ (p\geq 1)$ 和它的任意开球对等，即

$$l^p\sim B(x_0,\ \varepsilon)$$

证明　构造映射

$$F:B(x_0, \varepsilon) \to B(0,1),\quad \forall x \in B(x_0, \varepsilon), F(x) = \frac{x - x_0}{\varepsilon}$$

令

$$x = (x_1, x_2, \dots),\ x_0 = \left(x_0^{(1)}, x_0^{(2)}, \dots\right), x_i, x_0^{(i)} \in C, i = 1, 2, \cdots$$

且 $\sum_{k=1}^{\infty}|x_i|^p < \infty, \sum_{k=1}^{\infty}|x_0^i|^p < \infty$，则有

$$F(x) = F(x_1, x_2, \dots) = \left(\frac{x_1 - x_0^{(1)}}{\varepsilon}, \frac{x_2 - x_0^{(2)}}{\varepsilon}, \cdots\right)$$

F 是单射是显然的，下证 F 是满射。

$\forall y \in B(0, 1)$, $y = (y_1, y_2, \dots)$，$\|y\|_p = \left(\sum_{i=1}^{\infty}|y_i|^p\right)^{\frac{1}{p}} < 1$, 令 $u_i = \varepsilon y_i + x_0^{(i)}$，由 l^p 是线性空间得

$$u = \varepsilon y + x_0 \in l^p$$

又 $\forall y \in B(0,1) \Rightarrow \left(\sum_{i=1}^{n}\left|u_i - x_0^{(i)}\right|^p\right)^{\frac{1}{p}} = \left(\sum_{i=1}^{n}|\varepsilon y_i|^p\right)^{\frac{1}{p}} = \varepsilon\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}} < \varepsilon$,所以

$$u = \varepsilon y + x_0 = (u_1, u_2, \dots) \in B(x_0,\ \varepsilon),$$

可见 F 又是满射，因此 $B(0,\ 1) \sim B(x_0,\ \varepsilon)$。

构造映射

$$G: l^p \to B(0,\ 1),\quad \forall x = (x_1, x_2, \dots) \in l^p, x_i \in C,\ G(x) = G(x_1, x_2, \dots)$$

$$= (H(x_1), H(x_2), \dots) = \frac{x}{1 + \|x\|_p}, H(x_i) = x_i \Big/ \left(1 + \sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}}, i = 1, 2, \cdots$$

G 是单射是显然的，　下证 G 是满射

$\forall z \in B(0,\ 1), z = (z_1, z_2, \cdots), z_i \in C$,则有

$$\sum_{i=1}^{\infty}|z_i|^p < +\infty, \text{且} \|z\|_p = \left(\sum_{i=1}^{\infty}|z_i|^p\right)^{\frac{1}{p}} < 1$$

于是，令

$$y = (y_1, y_2, \cdots) = \left(\frac{z_1}{1 - \|z\|_p}, \frac{z_2}{1 - \|z\|_p}, \cdots\right) = \frac{z}{1 - \|z\|_p}, y_i = \frac{z_i}{1 - \|z\|_p}, i = 1, 2, \cdots$$

又

$$\sum_{i=1}^{\infty}|y_i|^p=\sum_{i=1}^{\infty}\frac{|z_i|^p}{\left(1-\|z\|_p\right)^p}=\sum_{i=1}^{\infty}|z_i|^p\Big/\left(1-\|z\|_p\right)^p<+\infty$$

因此 $y\in l^p$ 。因 $G(y)=\dfrac{z}{1-\|z\|_p}\Bigg/\left(1+\dfrac{\|z\|_p}{1-\|z\|_p}\right)=z$，所以 G 又是满射。故 $l^p\sim B$（x_0, ε）。

命题 9.1 和定理 9.10 的证法是类似的，关键在于构造两个双射，根据这两个双射，先求出其中一个的逆映射，再与另一个双射复合，得到一个新的双射，从而给出一种简捷的证法。下面采用这一证法证明命题 9.2 与定理 9.11。

命题 9.2 定理 9.8 中的赋范线性空间 $L^p[a, b]$ $(p\geq 1)$ 和它的任意开球对等，即

$$L^p[a, b]\sim B(f_0, \varepsilon)$$

证明 构造映射

$$T:L^p[a, b]\to B(f_0, \varepsilon),\forall f\in L^p[a, b],T(f)=f_0+\frac{\varepsilon f}{1+\|f\|_p}$$

下证 T 是双射。

显然 T 是单射，又 $\forall\, g\in B(f_0, \varepsilon)$，有

$$\|g-f_0\|_p=\left(\int_a^b|(g-f_0)(x)|^p\,\mathrm{d}x\right)^{\frac{1}{p}}<\varepsilon$$

即 ε-$\|g-f_0\|_p>0$，由于 $L^p[a, b]$ 是线性空间，$f_0, g\in B(f_0,\varepsilon)\subset L[a, b]$，所以

$$\frac{g-f_0}{\varepsilon-\|g-f_0\|_p}\in L^p[a, b]$$

令 $h=\dfrac{g-f_0}{\varepsilon-\|g-f_0\|_p}$，因

$$T(h)=f_0+\frac{\varepsilon h}{1+\|h\|_p}=f_0+\varepsilon\frac{g-f_0}{\varepsilon-\|g-f_0\|_p}\Bigg/\left(1+\frac{\|g-f_0\|_p}{\varepsilon-\|g-f_0\|_p}\right)=g$$

所以 T 又是满射，因此，$L^p[a, b]\sim B(f_0, \varepsilon)$。

抽去命题 9.1 和命题 9.2 中的具体范数，得到：

定理 9.11 若 X 是赋范线性空间，$B(x_0,\varepsilon)=\{x|\,\|x-x_0\|<\varepsilon, x\in X\}$ 是它的任意一个开球，则 $X\sim B(x_0,\varepsilon)$。

证明 构造映射

$$H:X\to B(x_0, \varepsilon),\quad H(x)=x_0+\frac{\varepsilon x}{1+\|x\|},\ \forall\, x\in X$$

显然 H 是单射，$\forall\, y\in B(x_0, \varepsilon)$，有 $y-x_0\in X$，且 $\|y-x_0\|<\varepsilon$，于是，存在

$$\frac{y-x_0}{\varepsilon-\|y-x_0\|}\in X$$

令 $u=\dfrac{y-x_0}{\varepsilon-\|y-x_0\|}$，因为

$$H(u)=H\left(\frac{y-x_0}{\varepsilon-\|y-x_0\|}\right)=x_0+\varepsilon\frac{y-x_0}{\varepsilon-\|y-x_0\|}\bigg/\left(1+\frac{\|y-x_0\|}{\varepsilon-\|y-x_0\|}\right)=y$$

所以 H 又是满射，这就证明了 $X\sim B(x_0,\varepsilon)$ 。

注 当 $H(x)=x_0+\dfrac{\varepsilon x}{1+\|x\|}$ 和 $H^{-1}(x)=\dfrac{x-x_0}{\varepsilon-\|x-x_0\|}$ 都连续时，$X\cong B(x_0,\varepsilon)$，即有

定理 9.12　设 X 是赋范线性空间，$B(x_0,\varepsilon)=\{x\,|\,\|x-x_0\|<\varepsilon, x\in X\}$ 是它的开球，若 X 和 $B(x_0,\varepsilon)$ 之间的对等映射是拓扑变换，则 $X\cong B(x_0,\varepsilon)$。

证明　将 X 和 $B(x_0,\varepsilon)$ 看作离散拓扑空间，由定义 9.1 和定义 9.2 立即得证。

在欧氏空间 R^n 中引进范数 $\|x\|=d(x,0), x\in R^n$，则有

定理 9.13　Banach 空间 R^n 与它的任意一个开球同胚，即 $R^n\cong B(x_0,\varepsilon)$。

证明　由定理 9.10 的证明，构造映射

$$L:R^n\to B(x_0,\varepsilon).\quad \forall x\in R^n,\ L(x)=x_0+\frac{\varepsilon x}{1+\|x\|}$$

L 是双射的证明同定理 9.11 的证法。下证 $L(x)=x_0+\dfrac{\varepsilon x}{1+\|x\|}, (x\in R^n)$ 是连续的。

设 $x=(x_1,x_2,\cdots,x_n)$，$x_0=\left(x_0^{(1)},x_0^{(2)},\cdots,x_0^{(n)}\right), x_i, x_0^{(i)}\in R, i=1,2,\cdots,n$，则

$$\begin{aligned}L(x)&=L(x_1,x_2,\cdots,x_n)\\&=(\phi_1(x_1,x_2,\cdots,x_n),\phi_2(x_1,x_2,\cdots,x_n),\cdots,\phi_n(x_1,x_2,\cdots,x_n))\end{aligned}$$

其中

$$\phi_i(x_1,x_2,\cdots,x_i,\cdots,x_n)=x_0^{(i)}+\frac{\varepsilon x_i}{1+\sqrt{\sum_{k=1}^{n}x_k^2}},\ i=1,2,\cdots,n$$

$\phi_i(x_1,x_2,\cdots,x_i,\cdots,x_n)$ 是关于 x_i 的一元初等函数，在 R 上连续，由定理 9.9，$L(x)=x_0+\dfrac{\varepsilon x}{1+\|x\|}, (x\in R^n)$ 连续，因此 $R^n\cong B(x_0,\varepsilon)$。

习题九

试建立区间 $[0,1]$ 与 $(0,1)$ 之间的对等映射。

参考文献

[1] 韩中庚. 数学建模方法及其应用 [M]. 北京：高等教育出版社，2010.

[2] 陈汝栋，于延荣. 数学模型与数学建模 [M]. 北京：国防工业出版社，2009.

[3] 段克峰. 输油管线建设费用的数学模型[J]. 陇东学院学报，2011，22(4)：7-10.

[4] 姜启源，谢金星. 数学建模案例选集 [M]. 北京：高等教育出版社，2006.

[5] 谢金星. 数学建模案例选集 [M]. 北京：高等教育出版社，2006.

[6] 王世儒，王金金，冯有前. 计算方法 [M]. 西安：西安电子科技大学出版社，2004.

[7] 姜启源. 数学模型 [M]. 北京：高等教育出版社，1987.

[8] 刘承平. 数学建模方法 [M]. 北京：高等教育出版社，2002.

[9] 王树禾. 数学模型选讲 [M]. 北京：科学出版社，2008.

[10] 杨启帆. 数学建模 [M]. 北京：高等教育出版社，2005.

[11] 东北师范大学微分方程教研室. 常微分方程 [M]. 2 版. 北京：高等教育出版社，2009.

[12] 王明新. 偏微分方程基本理论 [M]. 北京：科学出版社，2009.

[13] 段克峰. 基于一种复合模型的中国人口预测模型 [J]. 统计与决策，2012(20)：30-32.

[14] 段克峰. 基于 Malthus 模型的温室蔬菜杀虫剂合理使用方案 [J]. 安徽农业科学，2011，39(24)：14655-14657.

[15] 李庆扬. 数值分析 [M]. 5 版. 北京：清华大学出版社，2008.

[16] 周义仓，赫孝良. 数学建模实验 [M]. 西安：西安交通大学出版社，2001.

[17]汪国强. 数学建模优秀案例选编[M]. 广州：华南理工大学出版社，2001.

[18] 吴赣昌. 概率论与数理统计 [M]. 北京：中国人民大学出版社，2009.

[19] 冯杰，黄力伟，王勤，等. 数学建模原理与案例 [M]. 北京：科学出版社，2007.

[20] 赵静，但琦. 数学建模与数学实验 [M]. 北京：高等教育出版社，2000.

[21]宋世德，郭满才. 数学实验[M]. 合肥：中国科学技术大学出版社，2000.

[22] 朱道元，等. 数学建模案例精选 [M]. 北京：科学出版社，2005.

[23] 杨启帆，方道元. 数学建模 [M]. 杭州：浙江大学出版社，2005.

[24] 王树禾. 经济与管理科学的数学模型 [M]. 合肥：中国科学技术大学出版社，2000.

[25] 王红卫. 建模与仿真 [M]. 北京：科学出版社，2002.

[26] 李志林，欧宜贵. 数学建模及典型案例分析 [M]. 北京：化学工业出版社，2007.

[27] 姚金斌，李志林. 计算机模拟数学建模 [J]. 湖南理工学院学报：自然科学版，2008，21(1)：20-22.

[28] 王行仁. 建模与仿真的回顾及展望 [J]. 系统仿真学报，1999，11(5)：309-311.

[29] 陶利民，郭俊恩. 改进遗传算法在求解 TSP 问题上的应用研究 [J]. 计算机工程与应用，2009，45(33)：45-48.

[30]段克峰. 几种拓扑变换的构建及应用[J]. 宁夏师范学院学报，2009(3)：91-94.

[31] 段克峰，赵慧霞，姚晓斌，等. $(-\infty,+\infty)\cong(a,b)$ 的几种证法及推广 [J]. 甘肃联合大学学报：自然科学版，2009(6)：26-28.

[32] 段克峰. 赋范线性空间与其开球对等和同胚问题 [J]. 宁夏师范学院学报，2010(6)：10-13.

[33] 罗建军，等. MATLAB 教程 [M]. 北京：电子工业出版社，2005.

附录　数学建模训练题

1.一个公司在6座城市有分公司c_1, c_2, c_3, c_4, c_5, c_6，下面矩阵的i, j号元素是c_i到c_j的飞机票价，试为该公司制作一张由c_1到各公司去的最便宜的通航路线图。

$$\begin{bmatrix} 0 & 50 & \infty & 40 & 25 & 10 \\ 50 & 0 & 15 & 20 & \infty & 25 \\ \infty & 15 & 0 & 10 & 20 & \infty \\ 40 & 20 & 10 & 0 & 10 & 25 \\ 25 & \infty & 20 & 10 & 0 & 55 \\ 10 & 25 & \infty & 25 & 55 & 0 \end{bmatrix}$$

2.船工把一只老狼、一只羊和一筐大白菜运过河去，船小，一次只能运走一宗，为安全，不能让狼与羊，或羊与白菜在无人看管时在一起，试问如何运送最省时间又安全?

3.两人有酒，装满8斤之瓶，另有能装5斤与3斤的空瓶各一个，今欲平分其酒，请设计一个最简便的分法。

4.有一迷宫，其模型是K_6，但游人事先不知为K_6，试给出游人的遍历路线图。

5.a，b，c，d，e，f6人组成检查团，检查5个单位的工作，若某单位与检查团中的某一成员有过工作联系，则不能让他到该单位检查工作，已知第一单位与b，c，d有过联系，第二单位与a，e，f有过联系，第三单位与a，b，e，f有过联系，第四单位与a，b，d，f有过联系，第五单位与a，b，c有过联系，请列出去各单位检查人员的名单(要尽可能人多)。

6. $\{ace,be,dab,ab,be\}$这组信息是否可以分别把每个字用它的一个字母来代替，加以简化?若能，给出简化方案。

7.A，B，C，D4种材料制造a，b，c，d4种产品的成本如下面矩阵所示：

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} 99 & 6 & 59 & 73 \\ 79 & 15 & 93 & 87 \\ 67 & 93 & 13 & 81 \\ 16 & 79 & 86 & 26 \end{bmatrix} \end{array}$$

一种产品只用一种材料，问什么生产方案使成本最低？

8.用最少几种颜色可染中国地图，使邻省异色？

9.学校有7名教师，12个班，每周5天教学，教课要求由下面的矩阵给出

$$\begin{array}{cccccccccccc} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} & y_{11} & y_{12} \end{array}$$

$$P=\begin{bmatrix} 3 & 2 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 1 & 3 & 6 & 0 & 4 & 2 & 5 & 1 & 3 & 3 & 0 & 4 \\ 5 & 0 & 5 & 5 & 0 & 0 & 5 & 0 & 5 & 0 & 5 & 5 \\ 2 & 4 & 2 & 4 & 2 & 4 & 2 & 4 & 2 & 4 & 2 & 3 \\ 2 & 5 & 2 & 2 & 0 & 3 & 1 & 4 & 4 & 3 & 2 & 5 \\ 5 & 5 & 0 & 0 & 5 & 5 & 0 & 5 & 0 & 5 & 5 & 0 \\ 0 & 3 & 4 & 3 & 4 & 3 & 4 & 3 & 4 & 3 & 3 & 0 \end{bmatrix}\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{array}$$

其中 p_{ij} 是教师 x_i 教 y_i 班的节数，$P=(p_{ij})_{7\times 12}$，

（1） 一天应分几节课？

（2）若每天上8节课，需要几间教室？

10.比利时（d′ Hondt）分配方案：

将甲、乙、丙三系的人数都用1，2，3，…去除，将商从大到小排列，取前21个最大的，这21个中各系占有几个，就分给几个席位，你认为这种方法合理吗？

11.学校共有1000名学生，235人住在A楼，333人住在B楼，432人住在C楼. 学生们要组成一个10人委员会，使用 Q 值方法及 d′ Hondt 方法给出分配方案. 如果委员会为15人，分配方案是什么？

12.在凌晨1时警察发现一具尸体，测得尸体温度是29℃，当时环境的温度是21℃，1小时后尸体温度下降到27℃，若人体的正常温度是37℃，估计死者的死亡时间。

13.总统与首相面前同时送上同温度的热咖啡，总统在送到咖啡后立即加上一点冷奶油，等了10分钟才喝；首相则等了10分钟之后添加等量的冷奶油开始喝，问谁喝的咖啡热一些？

14.驱逐舰在浓雾中搜索潜水艇，雾一度散开，其时发现潜艇在3千米外的海面上，但潜艇立即下潜。驱逐舰速度2倍于潜艇，且知潜艇下潜后立即朝某未知方向直

线行驶。问驱逐舰应采取何种路线才能保证它直接开到潜水艇上方？

15.爱因斯坦狭义相对论中 $m=\dfrac{m_0}{\sqrt{1-\dfrac{v^2}{c^2}}}$，其中，$c$ 是光速，m 是运动粒子的质量，m_0是其静止质量，v 是质点运动速度。

（1）若质点在真空中从静止出发，在不变的引力场中做长时间运动，试求 $v(t)$，且证明 $\lim\limits_{t\to\infty}(t)=c$。

（2）令 $M=m-m_0$，若取其能量的相应增量 E 为其所受力 F 对它所做的功，从而

$$E=\int_0^v F\mathrm{d}x=\int_0^v \frac{\mathrm{d}}{\mathrm{d}t}(mv)\mathrm{d}x=\int_0^v v\mathrm{d}mv\ ,$$

试证 $E=Mc^2$。

16.若不考虑移民和高杀人率，纽约城的人口将满足方程

$$\frac{dp}{dt}=\frac{1}{25}p-\frac{1}{25\times10^6}p^2\ ,$$

其中，t 以年度量。

（1）事实上，每年有6000人从该城迁出，又有4000人被杀，试修正上面方程。

（2）已知1970年纽约城人口为800万，求未来任何时刻的人口，且求出 $t\to\infty$ 时的极限。

17.在什么年龄你的朋友们能最快地结婚？22岁？25岁？还是30岁？社会学家和心理学家们认为，在同等情况的人群中，存在着公开的或隐蔽的一种社会压力，例如“男大当婚，女大当嫁”等舆论，设 $m(t)$ 是 t 时刻某年龄段已婚人的比例，试建立此年龄段结婚速度的数学模型，且求解，并分析结婚速度的变化情况，指出如何确定未知参数。

18.请查看网络资源，建立我国综合国力的数学模型，并加以讨论。

19.工业废物注入盛有1000 cm^3的水池中，注入速度为1 $\mathrm{m}^3/\mathrm{min}$，充分搅拌后的混合物又以相同速度流入水池。

（1）求 t 时刻水池中废物的浓度。

（2）废物浓度达到20%，需要多长时间？

20.设 $t<0$时所有的车辆都一路纵队地以 v_0速度行驶，$t=0$时，前面的车突然刹车，求它后面的第一辆、第二辆车的运动速度 $v_1(t)$ 与 $v_2(t)$，讨论什么情况会撞车？

21.物种互助模型为

$$\begin{cases} \dfrac{\mathrm{d}p}{\mathrm{d}t} = -2x + 4y \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = x - 2y \end{cases},$$

（1） 当 t=0 时，x=100，y=300，试求 x（t）与 y（t）;

（2） 画出相图说明 $t \to +\infty$ 时会发生何现象。

[General Information]
书名=数学建模方法论
作者=段克峰主编；许万银副主编
页数=167
SS号=13439955
DX号=
出版日期=2013.08
出版社=兰州大学出版社